AD________________

COOPERATIVE AGREEMENT NO: DAMD17-95-2-5017

TITLE: Support of Study Entitled; "Conformationally Restricted Synthetic AIDS Vaccine"

PRINCIPAL INVESTIGATOR(S): Arnold C. Satterthwait, Ph.D.

CONTRACTING ORGANIZATION: Scripps Research Institute
La Jolla, California 92037

REPORT DATE: July 1996

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
July 1996

3. REPORT TYPE AND DATES COVERED
Annual (1 Jul 95 - 30 Jun 96)

4. TITLE AND SUBTITLE
Support of Study Entitled; "Conformationally Restricted Synthetic AIDS Vaccine"

5. FUNDING NUMBERS
DAMD17-95-2-5017

6. AUTHOR(S)
Arnold C. Satterthwait, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Scripps Research Institute
La Jolla, California 92037

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick
Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
HIV-1 initiates infection by attaching itself to CD4 on immune system cells. Two reagents, IgG1b12 and tetrameric CD4-IgG2, bind the envelope glycoprotein, gp120, on HIV-1, block this reaction and prevent infection. We have focused our research on the identification and mimicry of epitopes recognized by these reagents for use as a vaccine. We present 4 reports. A constrained DPE peptide loop based on Asp368ProGlu370 from the gp120 sequence has been used to isolate a neutralizing MAb similar to IgG1b12. A 3-D structure prediction for the V1/V2 domain on gp120 which plays a role in the neutralization of HIV-1 by IgG1b12 is presented and provides a basis for the design of constrained peptide vaccines. Both phage-peptide and synthetic peptide libraries have been extensively screened with Fab 12 and sCD4. Synthetic peptides characterized by Asp, Glu and rich in trptophans have been identified that block sCD4/gp120 and Fab 12/gp120 binding with IC50 = 10-100 uM. Finally, the DPE loop has been ligated to a "universal" T-cell epitope to yield a pure, well characterized four-branched loop-MAPS that is immunogenic in outbred mice and stimulates antibodies that bind recombinant gp120 IIIB.

14. SUBJECT TERMS
HIV-1, vaccine, peptides, CD4 binding site, gp120

15. NUMBER OF PAGES
76

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet *optical scanning requirements*.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	PR	- Project
G	- Grant	TA	- Task
PE	- Program Element	WU	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.
DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - Leave blank.
NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

ACS In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Arnold Sattler [signature] 9/3/96

PI - Signature Date

Table of Contents

PERSONNEL PAID BY US ARMY GRANT

		Date	% Effort
Arnold C.Satterthwait	P.I.	7/1/95 - 6/30/96	60%
Carlos F. Barbas, III	Co-P.I.	7/1/95 - 6/30/96	10%
Sara Venturini	Rsch Assoc	1/16/96 - 6/30/96	80%
EdelmiraCabezas	Rsch Assoc	12/1/95- 6/30/96	100%
Ruoheng Zhang	Rsch Assoc	5/1/96- 6/30/96	100%
Ming-zhu Zhang	Rsch Assoc	8/1/95 - 6/30/96	100%
Ellen Doval-Santos	Rsch Assoc	11/22/95- 6/30/96	100%
Weiping Hu	Rsch Assoc	7/1/95- 6/30/96	90%
Xu Yuan	Rsch Tech	4/22/96 -6/30/96	100%

Bibliography

1. Parren,W.H.I., Ditzel, H.J., Gulizia, R.J., Binley, J.M., Barbas III, C.F., Burton, D. R. and Mosier, D.E. (1995) Protection against HIV-1 infectioninhu-PBL-SCID mice by passive immunization with a neutralizing human monoclonal antibody against the gp120 CD4-binding site. AIDS 9:F1-F6.

2. Kolinski, A., Skolnick, J., Godzik, A. and Hu, W.P. (1996). A method for the prediction of surface "U"-turns and transglobular connections in small proteins. Proteins, in press

3. Satterthwait, A.C., Cabezas, E., Calvo, J.C. , Wu, J.X., Wang, P.L., Chen, S.Q., Kaslow, D.C., Livnah, O. and Stura, E.A., (1996) Constrained synthetic peptide vaccines, Peptides: Chemistry, Structure and Biology (Proceedings of the Fourteenth American Peptide Symposium), P.T.P. Kaumaya and Hodges, R.S. (Eds.) Mayflower Scientific Ltd. pp. 772--773

4. Cabezas, E., Wang, P.L. and Satterthwait, A. C., (1996). Identifying determinants of protein structure with loop peptides, Peptides: Chemistry, Structure and Biology (Proceedings of the Fourteenth American Peptide Symposium) P.T.P. Kaumaya and Hodges, R.S. (Eds.) Mayflower Scientific Ltd. pp. 734-735 (1996).

5. Cabezas, E., Stanfield, R., Wilson, I.A. and Satterthwait, A.C. (1996). Defining conformational requirements for the principle neutralizing epitope on HIV-1. Peptides: Chemistry, Structure and Biology (Proceedings of the Fourteenth American Peptide Symposium) P.T.P. Kaumaya and Hodges, R.S. (Eds.) Mayflower Scientific Ltd. pp. 800-801, (1996).

6. Ghiara, J. B., Ferguson, D. C., Satterthwait, A. C., Dyson, H. J. and Wilson, I. A., (1996). Structure-based design of a constrained-peptide mimic of the HIV-1 V3 loop neutralization site, Proc. Natl. Acad. Sci. , submitted

7. Satterthwait, A. C., Cabezas, E., Santos, E. D., Zhang, M.Z., Xie, Y. and Zhang, R., The conformational specificity of antibodies for peptides, Abstract for talk to International Chinese Peptide Symposium.

Selection of constrained loop libraries for binding to the b12 antibody family and a selection based study of a gp120 C3-region peptide with a antibody library prepared form an HIV-1 infected individual.

by Carlos Barbas III

Department of Molecular Biology, The Scripps Research Institute
10666 North Torrey Pines Road, La Jolla, CA 92037, USA

Introduction

Potential viral vaccines are often evaluated by immunization of animals and measurement of the functional activity of serum antibodies in vitro. This is followed by animal protection studies and finally trials in humans. Again the levels of neutralizing serum antibodies induced by immunization are used as a criterion of likely vaccine efficacy. Since human antibody responses, or versions thereof can now be cloned as bacteriophage display libraries, the possibility exists to challenge the human response in vitro for vaccine evaluation. Our studies have focused on the characterization of the broadly neutralizing antibody IgG12 (1-9) and evolved versions of this antibody which bind gp120 with higher affinity (10,11). With this antibody we have attempted to address both of the issue of epitope recognized and the ability of constrained peptides derived from gp120 to select neutralizing antibodies from antibody libraries prepared from HIV infected individuals.

Experimental Methods

Titering colony forming units (cfu). Phagemids which have been packaged into virions are capable of infecting male E. coli to form colonies on selective plates. Phage (packaged phagemid) was diluted in

SB (10^{-3}, 10^{-6}, and 10^{-8} fold dilutions) and 1μl was used to infect 50 μl of fresh OD 600 = 1 E. coli XLI-Blue grown in SB containing 10 μg/ml tetracycline. Phage and cells were incubated at room temperature for 15 min, then directly plated on LB/carbenicillin plates.

Amplification of Synthetic Antibody libraries. Synthetic libraries (12-14) were amplified by the addition of 10^{10} cfu of each library to 500 ml E. coli XLI-Blue grown in SB containing 10 μg/ml tetracycline. Following a 30min infection at 37C, 20 μg/ml carbenicillin was added and the culture shaken for 1hr at 37. The carbenicillin level was raised to 50 μg/ml and the culture shaken for 1hr at which time 10^{12} pfu of VCSM13 helper phage was added. Following 2 additional hours of shaking, 500ml of SB containing 50 μg/ml carbenicillin and 10 μg/ml tetracycline was added and kanamycin was then added to a final concentration of 70 μg/ml. The culture was grown 14 hrs at 37C before harvesting. The supernatant was cleared by centrifugation at 4°. Phage were precipitated by addition of 4% w/v polyethylene glycol 8000 and 3% w/v NaCl and incubated on ice for 30 minutes followed by centrifugation . Phage pellets were resuspended in 2 ml PBS (50 mM phosphate, pH 7.2, 150 mM NaCl) and utilized directly for panning.

Panning of the combinatorial library to select antigen binders. The panning procedure described is a modification of that originally described by Parmley and Smith (15). 4 wells of a microtiter plate (Costar #3690) were coated overnight at 4° with 25 μl of 12.5 μg/ml antigen in 0.1M bicarbonate pH 8.6. The wells were washed twice with water and blocked by completely filling the well with 1% w/v bovine serum albumin (BSA) in PBS and incubating the plate at 37° for 1 h. Blocking solution was shaken out and 50 μl of the phage library (typically 10^{11} cfu) were added to each well and the plate incubated for 2 h at 37°.

Phage were removed and the plate washed once with water. Each well was then washed 10 times with TBS/tween (50mM Tris HCl, pH 7.5, 150mM NaCl, 0.5% tween 20) over a period of 1 h at room temperature.

The plate was washed once more with distilled water and adherent phage were eluted by the addition of 50µl of elution buffer (0.1 M HCl, adjusted to pH 2.2 with solid glycine containing 1 mg/ml BSA) to each well and incubation at room temperature for 10 min. The elution buffer was pipetted up and down several times, removed and neutralized with 3µl 2M Tris base per 50 µl of elution buffer used. Eluted phage were used to infect 2 ml fresh (OD 600 = 1) E. coli XLI-Blue cells for 15 min. at room temperature after which 10 ml SB containing 20 µg/ml carbenicillin, 10 µg/ml tetracycline was added. Aliquots (20,1 and 0.1 µl) were removed for plating to determine the number of phage (packaged phagemids) which were eluted from the plate. The culture was shaken for 1 h at 37° after which it was added to 100 ml SB containing 50 µg/ml carbenicillin, 10 µg/ml tetracycline and shaken for 1h after which helper phage VCS-M13 (10^{12} pfu) were added and the culture shaken for an additional 2 hours. After this time, 70 µg/ml kanamycin was added and the culture incubated at 37° overnight. Phage preparation and further panning was repeated as described above.

Preparation of Soluble Fab. Phagemid DNA from positive clones was isolated and digested with Spe I and Nhe I. Digestion with these enzymes produces compatible cohesive ends. The 4.7 kb DNA fragment lacking the gIII portion was gel purified (0.6% agarose) and self-ligated. Transformation of E. coli XLI-Blue afforded the isolation of recombinants lacking the gIII fragment. Clones were examined for removal of the gIII fragment by Xho I/Xba I digestion which yielded a 1.6 kb fragment. Clones were grown in 15 ml SB containing 50 µg/ml carbenicillin and 20 mM $MgCl_2$ at 37° until OD 600 of 0.2 was achieved. Isopropyl-(beta)-D-thiogalactopyranoside, (IPTG), 1 mM, was added and the culture grown overnight at 37°. Cells were pelleted by centrifugation at 4000 rpm for 15 min. in a JA10 rotor at 4°. Cells were resuspended in 3 ml PBS containing 34 µg/ml phenylmethylsulfonylfluoride and lysed by sonication on ice (2-4 min. , 50% duty). The debris was pelleted by centrifugation at 14,000 rpm in a JA-20 rotor at 4° for 15 min. The supernatant was used directly for ELISA analysis and was stored at -20°.

ELISA analysis of Fab supernates. ELISA wells were coated with gp 120 exactly as above, washed 5 times with water, blocked in 100µl 1% BSA/PBS for 1h at 37° then reacted with 25µl Fab supernatants for 1h at 37°. After washing 10 times with water, 25 µl of a 1 in 1000 dilution of alkaline phosphatase-conjugated goat anti-human IgG $F(ab')_2$ (Pierce) were added and incubated for 1h at 37°. Following 10 washes with water, 50µl of p-nitrophenyl substrate were added and color development monitored at 405nm.

Preparation of purified Fab fragments. Bacterial cultures of the above clones were each grown in 1 liter of superbroth (for 1 liter: 10g of 3-(N-morpholino)propanesulphonic acid, 30g of tryptone, 20g of yeast extract, pH7.0) containing 0.5% glucose, tetracycline (10µg/ml) and carbenicillin (100µg/ml). The flasks were incubated, shaking, at 37C for 8 hours after which time 1mM isopropyl ß-D-thiogalactopyranoside was added to each culture. These were then incubated for a further 12 hours at 25C. The cells were collected by centrifugation (4000 rpm for 10 min at 4C) and the pellets resuspended in 10ml PBS containing 34 µg/ml phenylmethylsulfonylfluoride and 1.5% streptomycin sulphate. The suspension was subjected to three freeze-thaw cycles followed by centrifugation (17000 rpm for 30 min at 4C). The supernatants were collected and cleared by filtration through 0.2µm filters. Soluble Fab was purified by a single pass of the supernatants over an affinity column as follows. The column, consisting of goat anti-human F(ab')2 antibody (Pierce) linked to Gamma Bind G Sepharose (Pharmacia), was equilibrated in 3 column volumes of 87.2% phosphate buffer (0.1M sodium phosphate, dibasic; 0.5M sodium chloride), 12.8% citrate buffer (0.05M citric acid; 0.5M sodium chloride). The supernatants were loaded in the same buffer and the column was washed until the OD280 of the pass-through returned to a baseline level. The Fab was then eluted in 10.8% phosphate buffer, 89.2% citrate buffer and the collected fractions were neutralized with 1M Tris-HCl, pH 9.0 and concentrated to a final volume of approximately 1ml (typical concentration 100-800 µg/ml).

Results and Discussion

In earlier studies a large panel of mutant gp120 envelopes were utilized in a screen for binding to the human antibody b12 (4). These studies indicated that b12 binds a highly conformational epitope with determinants in the V2, C3, and C5 regions of gp120. The goal of our initial studies was to select constrained peptides from a random library with the hope that the selected constrained peptide would be homologous to regions in the envelope which constitute the epitope of this antibody. Since the antigen/antibody interface of evolved antibodies are significantly optimized over naturally occurring molecules, evolved antibodies should perform as optimized templates for the selection of mimetopes. The CD4-binding site antibody b4/12 recognizes a discontinuous epitope. The epitope recognized by this antibody has obvious implications for the design of a therapeutic vaccine since this antibody is so broadly and potently neutralizing (1-9). A subunit vaccine which incorporates such a mimetope could focus a response to this important epitope. For these studies we have utilized our collection of 10 synthetic antibody libraries which serve as a source of constrained loop libraries with over 10^9 independent clones (12-14). These libraries present random loops of various lengths and structural constraints. Selection of these libraries for binding the evolved antibody 3B3 (10) over 6 rounds of panning was performed. Vectors producing phage displayed antibodies were processed to remove the gIII fusion partner as in earlier studies described for production of soluble Fab. ELISA screening and subsequent nucleic acid sequencing revealed 6 unique Fab clones which bound the Fab 3B3. The sequences of the HCDR3 region of these Fabs are given below.

3B3ID-1	TTTDG
3B3ID-2	TSGDV
3B3ID-3	TGEDT
3B3ID-4	RGRDM
3B3ID-5	SGRSQFELRPQSAFDS
3B3ID-6	GFYGVPRLFFLGPRDW

HCDR3 Sequences of Fabs Selected to Bind 3B3

Unfortunately, the selected sequences cannot be definitively assigned to corresponding viral envelope sequences. Fabs 3B3ID 1-3 which are rich in amino acids with hydroxylated side-chains such as threonine and serine and the carboxylate bearing amino acids aspartate or glutamate. These sequences are analogous in some ways to sequences found on gp120, for example residues 49-51 TTT of the C1 region, 147-151 SSSGE of V1, 364-368 SSGGD of the C3 region. Of these regions only the C3 residue 368 has previousely been identified as a major binding determinate of b12 (4). The RD sequence found in 3B3ID-4 and 6 corresponds to residues 376-377 of the C5 region is known to be a determinant of binding in the b12 antibody (4). If these antibodies were true mimics of the antigen gp120 it is expected that sera from HIV infected individuals would bind these Fabs. Further characterization of selected Fabs revealed they were not reactive with polyclonal sera from the individual from which b12 was isolated even though they were reactive with 3B3. This together with the observation that selected HCDR3 regions do not bear significant sequence identity with gp120 suggests they are not valid candidates for vaccine study.

Recent studies have shown that antibodies to a wide range of viral pathogens can be isolated from antibody phage display libraries prepared from immune donors (16). Generally if a given donor has significant serum titer to a pathogen then high affinity monoclonal antibodies to that pathogen can be isolated from a library prepared from that donor. Some of the antibodies isolated have been shown to possess potent viral neutralizing ability in vitro, and anti-viral efficacy in vivo. The repertoire of antibodies selected from a library will depend on a number of factors including the composition of the library determined by the immune status of the donor and the antigen used for selection. We suggest that the ability of an antigen to select neutralizing antibodies from a human antibody library could be used as a measure of the likely quality of the antigen as a vaccine. We sought to

explore this possibility by selection of Fab libraries prepared from an asymptomatic HIV infected individual with a cyclic C3 peptide, [JKQSSGG**D**P**E**IVGZ]C-NH_2, which encompasses an epitope critical for CD4 and IgGb12 binding (Asp 368 and Glu 370, in bold)(17). The peptide was covalently coupled to maleimide activate bovine serum albumin for panning purposes. In order to explore this hypothesis, we utilized the well characterized immune library from which antibody b12 was isolated (3,4). Following 6 rounds of selection of this library for binding to the peptide-BSA conjugate, phage DNA was processed to allow for production of soluble Fab. Single clones were then characterized by ELISA for their ability to bind the peptide-BSA conjugate and rgp120 LAI. All clones bound both the peptide conjugate as well as rgp120 with higher signals for rgp120 binding. Nucleic acid sequencing of the selected clones revealed that all selected clones were of the b13 family of CD4-binding site antibodies in earlier studies characterized from this donor and isolated by selection with rgp120 (3,4). Consistent with the selection using the peptide, this family of antibodies is highly dependent on the sequence of gp120 at positions 386 and 370 which are included in the peptide used for selection. The more potent neutralizing antibody b12 was not selected even though it was present in the library.

Conclusions.

Selection of constrained peptide libraries for binding to the human Fab 3B3 (10), which is an in vitro improved version of the b12 antibody, did not result in peptides which significantly mimic features of gp120 to allow for them to be reactive with polyclonal immune sera. The selected loop regions of the antibodies likely bind in a mode analogous to an anti-id and do not truly mimic the antigen. Subsequent studies will focus on constraining the C3 region of gp120 in the context of a large phage display library. With respect to the C3 region, selections from immune libraries with our first constrained C3 peptide reveal that this epitope is a relevant target for incorporation into a vaccine since it proved sufficient for the selection of CD4-binding site directed antibodies from phage display libraries. This peptide,

however, did not preferentially select for the broadly reactive and potently neutralizing antibody b12 which was present in the library. Optimization of the conformation of this epitope will be necessary in order to present the structural features which will allow for the selection of more potent neutralizing antibodies libraries and the induction of a potent neutralizing response in animals.

References

1. Burton, D.R., Barbas III, C.F., Persson, M.A.A., Koenig, S., Chanock, R.M. and Lerner, R.A. (1991) A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. Proc. Natl. Acad. Sci. USA 88:10134-10137.

2. Barbas III, C.F., Bjorling, E., Chiodi, F., Dunlop, N., Cababa, D., Jones, T.M., Zebedee, S.L., Persson, M.A.A., Nara, P.L., Norrby, E. and Burton, D.R. (1992) Recombinant Human Fab fragments neutralize human type 1 immunodeficiency virus in vitro. Proc. Natl. Acad. Sci. USA, 89:9339-9343.

3. Barbas III, C.F., Collet, T.A., Roben, P., Binley, J., Amberg, W., Hoekstra, D., Cababa, D., Jones, T.M., Williamson, R.A., Pilkington, G.R., Haigwoods, N.L., Satterthwait, A.C., Sanz, I., and Burton, D.R. (1993) Molecular profile of an antibody response to HIV-1 as probed by combinatorial libraries. J. Mol. Biol. 230:812-823.

4. Roben, P., Moore, J.P., Thali, M., Sodroski, J., Barbas III, C.F. and Burton, D.R. (1994) Recognition Properties of a Panel of Human Recombinant Fab Fragments to the CD4 Binding Site of GP120 Showing Differing Ability to Neutralise Human Immunodeficiency Virus Type 1. J. of Virology 68:4821-4828.

5. Moore, J.P., Cao, Y., Qing, L., Sattentau, Q.J., Pyati, J., Koduri, R., Robinson, J., Barbas III, C.F., Burton, D.R., and Ho, D.D. (1995) Primary Isolates of Human Immunodeficiency Virus Type 1 are Relatively Resistant to Neutralization by Monoclonal Antibodies to gp120, and Their Neutralization is not Predicted by Studies with Monomeric gp120. J. of Virol. 69:101-109.

6. Burton, D.R., Pyati, J., Koduri, R., Sharp, S.J., Thornton, G.B., Parren, P.W.H.I., Sawyer, L.S.W., Hendry, M.R., Dunlop, N., Nara, P.L., Lamacchia, M., Garratty, E., Stiehm, E.R., Bryson, Y.J., Cao, Y., Moore, J.P., Ho, D.D., and Barbas III, C.F. (1994) Efficient Neutralization of Primary Isolates of HIV-1 by a Recombinant Human Monoclonal Antibody. Science 266:1024-1026.

7. Parren, P.W.H.I., Ditzel, H.J., Gulizia, R.J., Binley, J.M., Barbas III, C.F., Burton, D.R., and Mosier, D.E. (1995) Protection against HIV-1 infection in hu-PBL-SCID mice by passive immunization with a neutralizing human monoclonal antibody against the gp120 CD4-binding site. AIDS, 9: F1-6.

8. Trkola, A., Pomales, A.P., Yuan, H., Korber, B., Maddon, P.J., Allaway, G.P., Katinger, H., Barbas III, C.F., Burton, D.R., Ho, D.D., and Moore, J.P. (1995) Cross-Clade Neutralization of Primary Isolates of Human Immunodeficiency Virus Type I by Human Monoclonal Antibodies and Tetrameric CD4-IgG. J. of Virology, 69, 6609-6617.

9. Gauduin, M.-C., Allaway, G.P., Maddon, P.J., Barbas III, C.F., Burton, D.R., and Koup, R.A. (1996) Effective Ex Vivo Neutralization of Plasma HIV-1 by Recombinant Immunoglobulin Molecules. J. of Virol., 70, 2586-2592.

10. Barbas III, C.F., Hu, D., Dunlop, N., Sawyer, L., Cababa, D., Hendry, R.M., Nara, P.L. and Burton, D.R. (1994) In vitro evolution of a neutralizing human antibody to HIV-I to enhance affinity and broaden strain cross-reactivity. Proc. Natl. Acad. Sci. USA 91:3809-3813.

11. Yang, W.-P., Green, K., Pinz-Sweeney, S., Briones, A.T., Burton, D.R., and Barbas III, C.F. (1995) CDR Walking Mutagenesis for the Affinity Maturation of a Potent Human anti-HIV-1 Antibody into the Picomolar Range. J Mol. Biol. 254, 392-403.

12. Barbas III, C.F., Bain, J.D., Hoekstra, D.M. and Lerner, R.A. (1992) Semi-synthetic combinatorial antibody libraries: A chemical solution to the diversity problem. Proc. Natl. Acad. Sci. USA, 89:4457-4461.

13. Barbas III, C.F., Rosenblum, J.S. and Lerner, R.A. (1993) Direct selection of antibodies which coordinate metals from semisynthetic combinatorial libraries. *Proc. Natl. Acad. Sci. USA*, **14:**6385-6389.

14. Barbas III, C.F., Amberg, W., Somincsits, A., Jones, T.M. and Lerner, R.A. (1993) Selection of human anti-hapten antibodies from semisynthetic libraries. Gene, 137:57-62.

15. Parmley, S.F. and Smith, G.P. (1988) Gene 73, 305-318.

16. Burton, D.R., and Barbas III, C.F. (1994) Human antibodies from Combinatorial Libraries. Advances in Immunology 57:191-280.

17. The cyclic peptide was supplied by Arnold Satterthwait. See report on "The immunogenicity of a constrained peptide from the third constant region of HIV-1 gp120" by Arnold Satterthwait for the synthesis.

A 3-D structure prediction for the V1/ V2 domain of envelope glycoprotein 120 of HIV-1

by

Wei-Ping Hu, Angel R. Ortiz, Arnold C. Satterthwait and Jeffrey Skolnick

Department of Molecular Biology, The Scripps Research Institute,
10666 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

Introduction

There is a critical need for detailed structural information for gp 120, the exterior, envelope protein on HIV-1 that binds cellular CD4 and initiates infection. However, experimental determination of the 3-D structure of gp120 has been complicated by its large size and high sugar content. On the other hand, various epitope mapping studies have suggested structural features and imposed spatial constraints between sets of amino acid residues in this protein (1). This has led investigators to use the data for the construction of crude structural maps (2). However, sufficient data is now available to justify the use of computational methods for deriving low resolution structures.

As a first step, we explore data for the V1/V2 domain of gp120. We selected this domain for three major reasons. First, neutralizing epitopes have been mapped to the V2 region (3-6) and several of these epitopes overlap with amino acids required by IgG1b12 (7). A 3-D structural prediction for this region would provide a starting point for the design of conformationally restricted peptide vaccine candidates. Secondly, the V1/V2 domain have been fused to the C terminus of the N-terminal domain of gp70, the murine leukemia virus surface protein gp70 (8). As part of a chimeric protein, the V1/V2 domain binds conformationally dependent antibodies. Thus it is likely that the V1/V2 sequence provides complete information for the final fold. Thirdly, structure prediction is computationally intensive and at this stage best limited to domains of not more than about 100 residues.

The prediction of the three dimensional structure of a protein from its amino acid sequence is one of the most important, but unsolved, problems in contemporary molecular biology. There is no single approach that can predict all aspects of protein structure using only sequence information with absolute confidence. Thus, in this study we use a combination of several newly developed techniques to approach this task from different directions and combine the results to build a 3-D model for the V1/V2 region.

Although protein structure determination by experimental methods has become more efficient, the ratio between the number of known sequences and the number of known structures is rapidly increasing. An advantage of this fast growth of the protein sequence database is that many sequences can be grouped into structurally conserved families (9). This increasing number of protein families affords a new opportunity to exploit evolutionary information. When aligned together, such families exhibit features of residue conservation that are directly related to their three-dimensional structure. Using this principle, multiple aligned sequences in a family of homologous proteins have been used to improve the prediction of secondary structure in proteins (10).

Recently, a new algorithm for predicting surface "U" turn positions and the dominant secondary structures in transglobular linkers for small proteins has been reported by the Skolnick group (11). This method is based on a Monte Carlo search for the optimum division of the protein sequence into transglobular linkers and optimum assignments of structural templates to these linkers using a statistical potential. Very high accuracy is obtained in the predicted "U" turn positions (~95%), and the dominant secondary structure assignment is accurate in the range of 80%. This approach compliments traditional secondary structure predictions by predicting "U" turn positions for regions that have no well defined secondary structures which is critical for the molecular dynamics simulation for protein structures.

Prediction of residue contact maps has also been attempted by analyzing correlated mutations in multiple sequence analysis with interesting results (12). This approach builds on the idea that contact residues can be identified by noting compensating mutations in pairs or multiples of amino acids distant in sequence. Also, structure calculations based on multiple sequence alignments are starting to emerge and

are aimed at predicting simple protein topologies (13-15). However, the studies reported so far have not succeeded for *de novo* predictions. There are several reasons for the limited success of such approaches. First of all, the number of restraints required for a successful structure calculation using distance geometry or Monte Carlo simulations with an all-atom force field is substantial and is usually higher than the number of restraints that can be extracted from the multiple sequence alignment. Secondly, the restraints that need to be satisfied are of rather low accuracy and precision.

We recently proposed (16) a method that is effectively able to predict the low resolution structure of small proteins using restraints derived from multiple sequence alignment and Monte Carlo simulations. This method partially solves the problems mentioned above by using several new features. First, the signal-to-noise ratio of the predicted contacts is significantly improved using a new multiple-step procedure (17). In this procedure, one first identifies the topological elements of the protein using the "U" turn and transglobular linker prediction along with Rost and Sander's secondary structure prediction. Second, correlated mutation analysis is used to determine the pairs of topological elements that come in contact. Third, one expands the number of contacts with an inverse folding algorithm. Finally, a lattice Monte Carlo simulation, the MONSTER method (18), is used to generate the protein model. The MONSTER method can efficiently find native-like structures with a RMSD of 3-4 Å using N/5 contact map restraints, where N is the number of amino acid residues in the chain. This approach has been applied to a small set of topologically different proteins, and preliminary results have been promising (16).

In the current study, we use the combined approach outlined in the previous paragraph to build a 3-D structure for the V1/V2 domain of gp120 that is consistent with most of the structural data obtained from various methods involved. The model structure is then discussed in terms of antibody binding studies.

II. Calculations and Results

II.A. Sequence selection. There are many different HIV gp120 proteins available in the Los Alamos HIV sequence database. They have been grouped on the basis of sequence similarities into Clades. In order to improve the reliability of the predictions,

we selected V1/V2 domains from different Clades. To minimize the computational time, we further selected sequences with short hypervariable regions. There is no clear boundary for the V1/V2 region in the gp 120 sequence. However, there are three nested disulfide links in this region. The outermost disulfide link provides a natural boundary for this region since most antibody activities in the V1/V2 domain are retained on transplantation of this region to another protein (9). The "U" turn and secondary structure algorithm was applied to the V1/V2 domain from six sequences. These sequences are from HIVHXB2R (Clade B), HIVD687 (Clade A), HIVTB132 (Clade B), HIVNDK (Clade D), HIVMVP5180 (Clade E), and HIVANT70 (Clade E)(Table 1). The V1/V2 sequence from HIVHXB2R was studied further with a correlated mutation analysis, an inverse folding algorithm, and a lattice Monte Carlo folding simulation.

II.B. U-turn positions and secondary structures. The predictions from The "U" turn and secondary structure algorithm are summarized in Table I. There are 5 or 6 strongly predicted surface "U" turn positions which imply 6-7 topological elements (transglobular linkers). Both methods predict that these elements are mostly extended structures. However, the innermost disulfide link contains a "hypervariable" region signaling the presence of a large surface loop. This agrees with most predictions except those for HIVNDK and HIVANT70 which predict an alpha helix. In addition, the middle linker beginning about Q170 is predicted to be alpha helical for all sequences except that for HIVD687. Immediately following the predicted helix is a second hypervariable region which is predicted to be a loop. These predictions suggest that the V1/V2 domain adapts the following secondary structures in its native state: extended-(extended)-long surface loop-extended-helix or extended-short surface loop-(extended)-extended. Brackets indicate the addition of a short extended form in some sequences. The specific structure assignment for the HIVHXB2R sequence fragment is shown in Table II.

II.C. Correlated mutation analysis. Fifty HIV and SIV V1/V2 sequences were aligned with HIVHXB2R and used for correlated mutation analysis. This provides a set of potential contact residues for the V1/V2 residues that can be used for predicting a HIVHXB2R structure. Three screens were used to select contact residues from this set for further study. First, the contact must be long range, residing in different topological elements identified by our "U turn" and transglobular linker prediction. Second, the correlation coefficient of the possible contact must be greater than 0.85. Finally, for any

given pair of topological elements, the contact with the highest correlation coefficient was selected. The predicted contacts that satisfy these criteria and the corresponding contacting structural elements are listed in Table III.

II.D. Inverse Folding. After predicting the elements that are in contact, we expanded the number of contacts using an inverse folding algorithm (19) before performing lattice Monte Carlo simulation. We were concerned, however, that the long surface loop region from residue 14 to 38 might interfere with the tertiary structure formation during the simulation. Assuming that this loop region is not essential for folding, we decided to eliminate the 21 residue fragments from leucine 134 to isoleucine 154 for the inverse folding calculation. The structure assignments in Table II are converted for this simplified sequence in Table IV.

We then threaded overlapping 15 and 20 residue fragments from the simplified sequence through the structure database (19). We identified three 15-residue and three 20-residue fragments that gave high scores for related structural motifs. These fragments corresponded to elements 2-3, 5-6, and 6-7 in the sequence which are predicted to be in contact from the correlated mutation analysis. We then extracted contact patterns for the the simplified HIVHXB2R sequence on the basis of the high scoring motifs. The contacts generated by this procedure are listed in Table V.

II.E. Monte Carlo Simulation.

The structure assembly of the protein model is carried out using a lattice representation of the Cα-backbone trace. The details of the model have been described previously (20). Here, we give a brief summary for the reader's convenience.

Lattice model of protein chain

The Cα backbone is a string of vectors of the type $a \bullet \mathbf{v}$ with $\{\mathbf{v}\}=\{(3,1,1),\ldots (3,1,0),\ldots (3,0,0],\ldots (2,2,1),\ldots (2,2,0),\ldots.\}$. The value of the parameter a=1.22 Å has been optimized to obtain the best fit of the lattice representation to high resolution, sequentially non homologous protein structures from the Brookhaven Protein Data Bank,

PDB. The virtual bond angles between successive Cαs are restricted to reproduce a protein-like distribution. There is a single interaction center for each side chain group.

Force field of the protein model:

The force field contains potentials of mean that account for the short range interactions, long range interactions, and hydrogen bond interactions (which could be short or long range). All contributions to the potential are available via anonymous ftp (21). The total energy of the system is given by:

$$E= 0.5E_{14} + 1.5E_{1} + 2.75E_{pair} + E_{H\text{-}bond} + E_{target,14} + V_{long} \qquad (1)$$

- Short range interactions: This component depends on the sequence through the pair of neighboring amino acids A_i and A_{i+1} , and controls the local chain geometry. It considers six conformational states that roughly correspond to extended right, and extended left handed states, wide right and left turns, and right and left handed helical states respectively.

- Long range interactions: Encoded *via* two potentials of mean force. The first term is a one-body, centrosymmetric potential which reflects the tendency of some amino acids to be buried and some to be exposed. The second term corresponds to the pairwise tertiary interactions. These are neglected between nearest neighbors down the chain, since these short range interactions are already accounted for by the hydrogen bond potential (see below) and secondary structure preferences. The pair interactions beyond the fourth neighbor are derived from the statistics of the database. The amino acid pairwise specific parameters are described elsewhere (22).

- Hydrogen bond interactions: The model hydrogen bond potential operates only between α-carbons. Each α-carbon can participate in at most two hydrogen bonds (the α-carbon of proline is an exception and can participate in only one hydrogen bond), and there is no directionality (donor-acceptor) in the scheme. The scheme used reproduces most (about 90 %) of the main chain hydrogen bonds assigned to the structure by DSSP Kabsch-Sander algorithm (23). An explicit cooperativity is also introduced into the hydrogen bond scheme. When two neighboring pairs of residues are hydrogen bonded, then the system gets an additional favorable energy (cooperativity). Note that we ignore all side chain -side chain and side chain - backbone hydrogen bonding.

Restraint Contributions

Secondary structure restraints and a limited number of tertiary restraints are used. Furthermore, a set of knowledge-based restraints is used. The implementation of each kind of restraint is discussed in turn.

Short range restraints

1. For those residues assigned to be helical, hydrogen bonds beyond the fifth residue along the chain are not allowed. Similarly, α and β-assigned residue cannot have a helix hydrogen bond pattern nor can it hydrogen bond to a residue that has been assigned to be in a helical region of the molecule.

2. For left handed helical states, all residues experience a repulsion of 1.0, in k_BT units. On the other hand, turns are encoded on a generic basis, i.e., their chirality is not specified. Rather they behave as flexible joints between regular secondary structural elements.

Long range restraints

Long range restraints operate on the level of distances between the centers of mass of the side groups. If residues i and j are predicted to be in contact, then the residue based pair potential is modified so that e_{ij}=-1.25. The long range tertiary restraints are as follows. Let $d_{ij}=r_{ij}-R^{con}_{ij}$

$$
\begin{aligned}
Y_{long}(r_{ij}) &= 0 && \text{if } r_{ij} < R^{con}_{ij} \\
&= g\, d_{ij}^{\,2} && \text{If } R^{con}_{ij} < d_{ij} < 34.5\ \text{Å} \\
&= g(34.5)^2 && \text{otherwise.}
\end{aligned}
\qquad (2)
$$

Typically, the value of g ranges from 0.5 to 2.

In folding from random compact states, the restraints are not implemented simultaneously. Rather, the "sequential growing strategy" is used. This appears to increase the folding efficiency by decreasing the extent of kinetic trapping.

Conformational sampling

The sampling of conformational space occurs via a standard asymmetric Monte Carlo Metropolis scheme (24). The conformational updates are composed of several types of local conformational micromodifications of the chain backbone and their associated side groups, side group equilibration cycles, and rare (small distance) motions of larger chain fragments.

A fully extended chain is selected as initial structure in the assembly stage. The simulation is started at a reduced temperature in the range of 2.0, and then the temperature is slowly lowered to 1.0. Eight simulations starting from extended conformations were carried out. The predicted structure is the one which exhibits the lowest average and minimum energies.

II.F. All-atom reconstruction.

After a conformation is selected from the lattice Monte Carlo simulation, an all atom model building is carried out in order to generate an all atom structure of the molecule. This is done using the program MODELLER (25) and the structure has been further transformed by the AVS program for display in Figure 2.

III. Results and discussion

III.A. Structure description.

The predicted structure of the V1/V2 domain of gp120 (Figure 2) corresponds to an alpha/beta protein. In the predicted structure there are five strands that form a beta sheet type structure. There is one alpha helix predicted that packs against the beta sheet structure. The overall fold is reminiscent of the Protein G fold. Energetic characteristics of the conformation can be found in Table VI. From the total number of 25 restraints, 17 are satisfied in the final conformation (Table VII). The fraction of restraint satisfaction is similar to that obtained in previous cases in the development of the methodology (16). The final structure satisfies the experimental disulfide bridge restraints as well. However, at this stage it presents certain distortions in the first beta hairpin of the domain, probably

as a result of the loop deletion performed in that region. This suggests that elimination of the loop was a too drastic procedure and that probably it is necessary to include it in the simulations in order to correct the defects in that part of the structure.

III. B. Antibody binding.

Residues 176-177FY, 179/180LD, 183/184 PI and 191-193 YKL (Figure 2) are associated with conformational epitopes bound by neutralizing MAbs BAT-085(2,3), G3-4(2), G3-136(2), CRA-3(2) and 697-D (5). The affinity of BAT-085 for gp120 is enhanced 3-4 fold by mutations at 183/184 (PI to SG) and 191-193 (YSL to GGS) while the same mutations block binding to the other antibodies. Since BAT-085 binds the linear peptide 169-183 while the other MAbs require native gp120 (2,5), it is possible that enhancing mutations at 183/184 PI and 191-193 YKL play a conformational role. These mutations could either relieve a steric block to BAT-085 thus enhancing affinity or they could disrupt a local or global conformation freeing 176-177FY and 179/180LD to bind BAT085. Similarly, 183/184 PI and 191-193 YKL may make up part of the epitope for MAbs G3-4, G3-136 and 697-D. However, mutations at these amino acids could disrupt binding by an indirect conformational effect on 176-177FY and 179/180LD. In support of this latter interpretation, MAbs G3-4, G3-136 and 697-D bind weakly to peptides from the 161-181 region (2,5). Since 183/184 PI and 191-193 YKL are not included in these peptides, it might be argued that their effect on binding gp120 is conformational.

The calculated structure place PI and YKL on the opposite sides of a β-hairpin loop. It is reasonable that mutations could disrupt the conformation of this loop that in turn could influence the conformation of FY/LD which occupies a preceding partially exposed loop. Modification of FY completely abrogates affinity of IgG1b12 for gp120. The FY/LD epitope in the V2 region should be explored as a potential constrained peptide vaccine.

References

1. Moore, J. P. and Sodroski, J.(1996). Antibody cross-competition analysis of the human immunodeficiency virus type 1 gp120 exterior envelop glycoprotein, J. Vir. 70:1863

2. Moore, J. P. et al. (1993). Probing the structure of the V2 domain of human immunodeficiency virus type 1 surface glycoprotein gp120 with a panel of eight monoclonal antibodies: human immune response to the V1 and V2 domains, J. Vir. 67: 6136

3. Fung, M.S.C. et al. (1992). Identification and characterization of a neutralization site within the second variable region of human immunodeficiency virus type 1 gp120, J. Vir. 66:848 (1992)

4. McKeating et al. (1993). Characterization of neutralizing monoclonal antibodies to linear and conformation-dependent epitopes within the first and second variable domains of human immunodeficiency virus type 1 gp120, J. Virol. 67: 4932

5. Gorny, M.K. et al. (1994). Human anti-V2 monoclonal antibody that neutralizes primary but not laboratory isolates of human immunodeficiency virus type 1, J. Virol. 68:8312.

6. Ditzel, H. J. et al., (1995). Neutralizing recombinant human antibodies to a conformational V2- and CD4-binding site-sensitive epitope of HIV-1 gp120 isolated by using an epitope-masking procedure, J. Imm. 154: 893

7. Binding of IgG1b12 to gp120 is knocked out by mutations at **YKL** and diminished by mutations at **PI**. However, deletion of the full V1/V2 domain leads to only partial reduction in affinity. Carlos Barbas, personal communication.

8. Kayman, S. C. et al.(1994). Presentation of native epitopes in the V1/V2 and V3 regions of human immunodeficiency virus type 1 gp120 by fusion glycoproteins containing isolated gp120 domains, J. Vir. 68:400

9. Sander, C. & Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins 9, 56-68.

10. Rost, B. & Sander, C. (1993). Prediction of secondary structure at better than 70% accuracy. J. Mol. Biol. 232, 584-599.

11. Kolinski, A., Skolnick, J., Godzik, A. and Hu, W.-P. (1996). A method for the prediction of surface "U"-turns and transglobular connections in small proteins. Proteins, in press.

12. Gobel, U., Sander, C., Schneider, R. & Valencia, A. (1994). Correlated mutations and residue contacts in proteins. Proteins, 18, 309-317.

13. Hanggi, G. & Braun, W. (1994). Pattern recognition and self-correcting distance geometry calculations applied to myohemerythrin. FEBS Letters 344, 147-153.

14. Numenthaler, C. & Braun, W. (1995). Predicting the helix packing of globular proteins by self-correcting distance geometry. Prot. Science 4, 863-871.

15. Aszodi, A., Gradwell, M. J. & Taylor, W. R. (1995). Global fold determination from a small number of distance restraints. J. Mol. Biol. 248, 308-326.

16. Ortiz, A.R., Hu, W.-P., Kolinski,,A. and Skolnick, J., Method for the low resolution prediction of small protein tertiary structures (1996). submitted to the 2nd Pacific Symposium of BioComputing in Hawaii.

17. Ortiz, A. R., Hu, W. P. & Skolnick, J. (1996). Contact map prediction using correlated mutations and inverse folding, unpublished results

18. Skolnick, J., Kolinski, A. & Ortiz, A. R. (1996). MONSTER: A method for folding globular proteins with a small number of distance restraints. J. Mol. Biol.. (submitted).

19. Hu, W.-P., Godzik, A., Skolnick, J. (1996). Sequence-Structure Specificity-How Does an Inverse Folding Approach Work? Prot. Engn., submitted.

20. Kolinski, A. & Skolnick, J. (1996). *Lattice Models of Protein Folding, Dynamics and Thermodynamics*, R. G. Landes Company, Austin, TX.

21. Skolnick, J. (1996). ftp.scripps.edu in directory /pub/skolnick/nmr.

22. Skolnick, J., Jaroszewski, L., Kolinski, A. & Godzik, A. (1996). Derivation and testing of pair potentials for protein folding. Is the quasichemical approximation valid ?. Prot. Engn., submitted.

23. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577-2637.

24. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). Equation of state calculations by fast computing machines. J.Chem.Phys. 51, 1087-92.

```
HIV HXB2R   (CLADE B)
RES. #      115        125        135            145        155        165        175        185                  195        205
SEQUENCE    SLKPCVKLTP LCVSLKCTDL K....NDTNTNSSS GRMIMEKGEI KNCSFNISTS IRGKVQKEYA FFYKLDIIPI D........NDT..TSYKLT SCNTSVITQA CPKVS
"U" turn    EEEEEEEETT TTEETTTTTT T....TTTEEEEEE EEEEEETTTT TTTTEEEEET TEEEEHHHHH HHHHEEEEEE E........ETT..TTEEEE EEEETTEEEE EEEEE
PHD         LLLL???LLL ?EEEEEE?LL L....LLLLLLLLL LL???????? ?E?EE????? ????HHHH?? ?EE????EEE ?........LLL..L?EEEE ??LL?EE??? ??LLL

HIVD687     (CLADE A)
SEQUENCE        CVKLTP LCVTLNCSDV T....NNTFSNVS. ...YDMKGEI KNCSFNMTTE LRDKKRKIFT .FYRLDVVQI D........NSS..SEYRLI NCNTSAITQA C
"U" turn        EEEEEE EETTTTTTTT T....TTTTTTTT. ...TTEEEEE TTTEEEEEET TTTTTTTEEE .EEETTTEEE E........TTT..TEEEEE EEEEEEEEEE E
PHD         LLLL???LLL ?EEEEE??LL L....LLLLLLLL. ...LLLL??E EEEEE????? ??LLL????E .EEEEEEEEE ?........LLL..L?EEEE E?LL?EE??? ??L?L

HIVTB132    (CLADE B)
SEQUENCE        CVKLTP LCVTLNCTNL R....NDT...... .......SEM KNCSFKITTS IRDKLQKEYA LFYKLDIVPI NND......NTS....YRLI SCNTSIITQA C
"U" turn        EEEEEE ETTTTTTTEE E....EET...... .......TTT TTEEEEEETT TTHHHHHHHH HHHHTTTTEE EET......TTT....TEEE EEEEEEEEEE E
PHD         LLLL???LLL ??EE?E??LL L....LLL...... .......L?? ?EEE?L???? ??L??HHH?? ??E????E?? LLL......LL?....EEEE E?LL?????? ??LLL

HIVNDK      (CLADE D)
SEQUENCE        CVKLTP LCVTLNCTDE LR...NSKG.NGKV ....EEEEKR KNCSFNV... .RDKREQVYA LFYKLDIVPI DNNNRT...NST...NYRLI NCDTSTITQA C
"U" turn        EEEEEE EETTTTTTTT TT...TTTT.HHHH ....HHHHTT TTTEEEE... .ETTTTTHHH HHHHHHHHHH HTTTTT...TTT...EEEEE EEEEEEEEEE E
PHD         LLLL???LL? ?EEEEEE?LL LL...LLLL.LLLL ....LLL??? ???EE??... .LL??????E EEE????EE. ?LLLLL???LLL???EEEEE E?LL?EE??? ???LL

HIVMVP5180  (CLADE E)
SEQUENCE        CEKMTF LCVQMNCVDL QTNKTGLL..NETI N.......EM RNCSFNVTTV LTDKKEQKQA LFYVSDLSKV ...NDSNAVNGT...TYMLT NCNSTIIKQA C
"U" turn        EEEEEE EEEETTTTTT TTTTTTTE..EETT T.......TT TTTTTTTEEE EETTTTTTTH HHHHHHHHTT ...TTTTTTTTT...TEEEE TTTTTEEEEE E
PHD         LLLL?E?LL? ?EEEEEE?LL LLLLLLLL..LLLL ?.......?? EEEEE????? ?????HH??? EEEE???E?? ...?LLLLLLL?...EEEEE E?LL?EEE?? ????L

HIVANT 70   (CLADE E)
SEQUENCE        CVQMTF LCVQMECTNI A....GTT...... .....NENLM KKCEFNVTTV IKDKKEKKQA LFYVSDLMEL ...NETSSTNKTNSKMYTLT NCNSTTITQA C
"U" turn        EEEEEE EEETTTTTTH H....HHH...... .....HHHHH HTTTTTTTEE EEETTTTTHH HHHHHHHHHH ...HHHEEEEEETTTEEEET TTTTTEEEEE E
PHD         LLL??E?LL? ??EEEEE?LL L....LLL...... .....LL??E EEEEE????? ?????HH??? EEEEE?EE?? ...?LLLLLLLLLL?EEEE E?LL?EE??? ????L
```

Table I. Structure predictions for the V1/V2 domain. Sequences are aligned to HXB2R. Each sequence is treated with the "U"-Turn/Linker and the PHD algorithms. E = extended, T = turn/loop and H = helix for the "U"-Turn/Linker algorithm. E = extended, L = loop, H = helix and ? = no strong prediction for the PHD algorithm.

Table II. Structure Assignment for the HIVHXB2R Sequence Fragment

Element No.	Residue Range[a]	Assignment
1	1–5	Extended
2	8–13	Extended
3	39–45	Extended
4	51–58	Helix/Extended
5	63–68	Extended/Loop
6	72–78	Extended
7	81–87	Extended

[a]Residues are numbered from the beginning of the V1/V2 domain, i.e. cysteine 119 is residue 1. The residues between the assigned elements are "U"-turns.

Table III. Predicted Contacts from Correlated Mutation Analysis for HIVHXB2R

Correlation Coefficient	Residue No.1	Residue No.2	Elements in Contact
1.00	11 (L)[a]	85 (Q)	2,7
0.91	66 (I)	85 (Q)	5,7
0.91	11 (L)	66 (I)	2,5
0.90	45 (T)	85 (Q)	3,7
0.90	11 (L)	45 (T)	2,3
0.86	40 (S)	66 (I)	3,5
0.84	75 (L)	85 (Q)	6,7
0.84	66 (I)	75 (L)	5,6
0.82	40 (S)	73 (Y)	3,6

[a]The letter in the parentheses represents the one-letter code for the amino acid at that position.

Table IV. Structure Assignment for the Simplified HIVHXB2R Sequence Fragment

Element No.	Residue Range	Assignment
1	1–5	Extended
2	8–13	Extended
3	18–24	Extended
4	30–37	Helix/Extended
5	42–47	Extended/Loop
6	51–57	Extended
7	60–66	Extended

Table V. Lists of the Contacts Generated from the Fragment Based Inverse Folding

No.	Residue[a] Pair	No.	Residue Pair	No.	Residue Pair	No.	Residue Pair
1	6,21	11	11,20	21	43,50	31	46,51
2	7,21	12	12,17	22	43,53	32	46,53
3	7,22	13	12,19	23	43,56	33	47,52
4	8,19	14	13,18	24	44,52	34	54,62
5	8,21	15	13,20	25	44,54	35	55,63
6	9,22	16	20,26	26	44,55	36	56,62
7	10,17	17	39,56	27	45,50		
8	10,19	18	40,56	28	45,52		
9	10,21	19	41,53	29	45,54		
10	11,18	20	42,54	30	45,56		

[a]The residue number is based on the simplfied HIVHXB2R sequence.

Table VI. Dissection of the different energetic contributions to the final model. Energies are expressed in kT units.

Restraint contribution	5.2
Knowledge-based rules penalty	1.7
Hydrophobic moment contribution	-4.3
Hydrogen bond contribution	-83.9
Hydrogen bond strand cooperativity	-5.2
Local entropy contribution	-91.0
Long range contribution	-115.9
TOTAL ENERGY	**-295.2**

Table VII. Detailed description of the restraint contributions to the total penalty energy. For each of the restraints, four columns are displayed. The fisrt column is the distance (in Å) between the center of mass of the corresponding side chains that are predicted to be in contact; in the second column the energy penalty (in kT units) in the final conformation is shown; in the third column the residues involved in the restraint are shown; and in the last column the "status" of the restraint, that is, the consideration of "satisfaction" or "no satisfaction" is shown.

SC-SC DIS	ENERGY	RESIDUES	STATUS
10.26136	.00000	4 63	---------
4.45928	.00000	6 61	SATISFIED
3.54647	.00000	8 57	SATISFIED
6.14859	.00000	10 21	SATISFIED
10.74796	.50135	11 25	---------
6.87888	.00778	12 19	---------
4.38512	.00000	13 18	SATISFIED
6.69208	.00000	13 20	SATISFIED
9.87949	.28227	19 55	---------
4.44431	.00000	20 26	SATISFIED
6.58523	.00000	24 64	SATISFIED
10.15782	.00000	39 56	---------
6.30780	.00000	40 56	SATISFIED
4.89462	.00000	41 53	SATISFIED
7.27518	.00000	42 54	SATISFIED
7.14999	.00000	43 50	SATISFIED
7.84463	.00000	43 53	SATISFIED
6.34302	.00000	44 52	SATISFIED
8.36638	3.47400	44 54	---------
5.16727	.00000	45 50	SATISFIED
5.04756	.00000	46 51	SATISFIED
4.36360	.00000	54 62	SATISFIED
3.50797	.00000	55 63	SATISFIED
7.24514	.81134	55 64	---------
7.34383	.12954	56 62	---------

Constraints= 25 # Satisfied = 17
Contacts |i-j|>1 ... 140

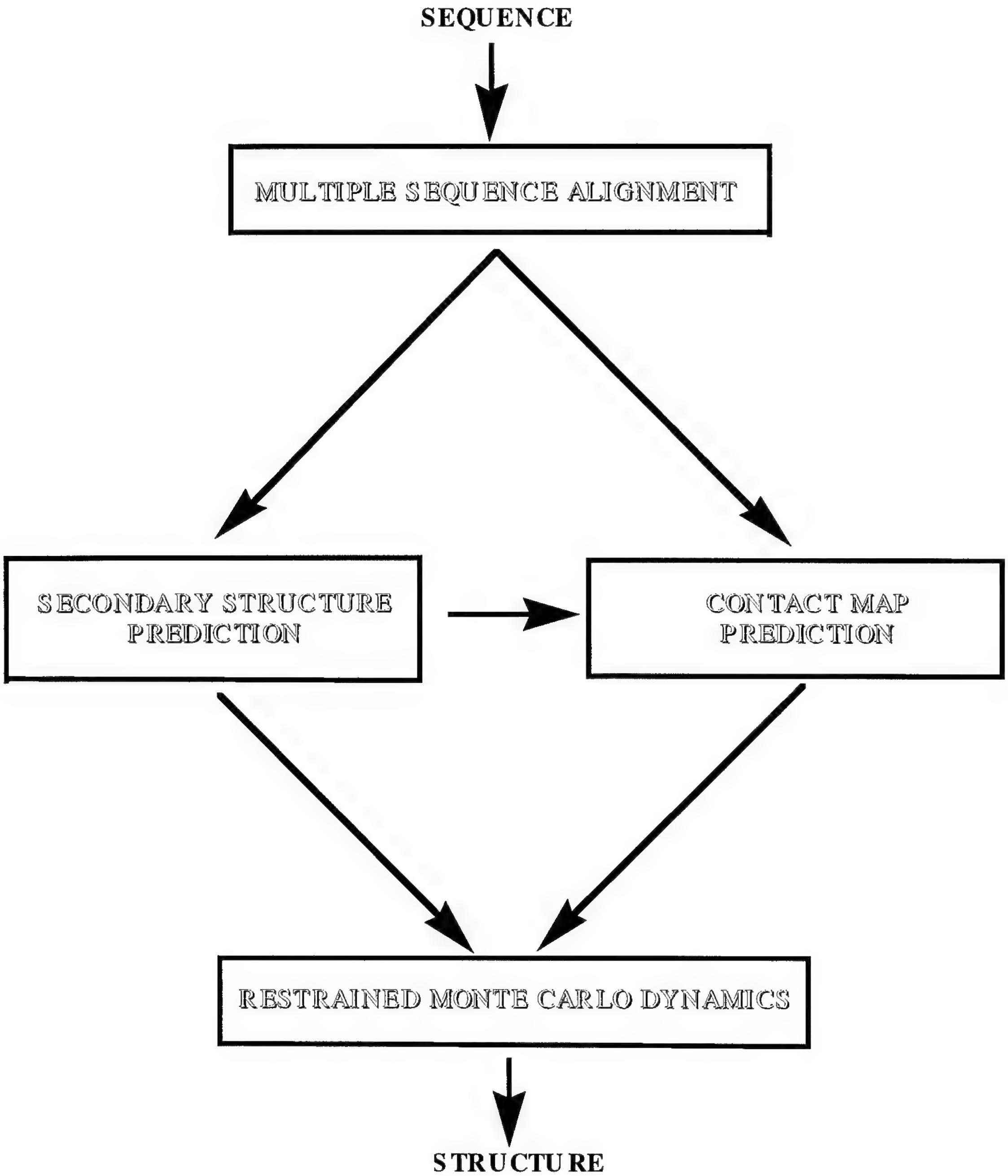
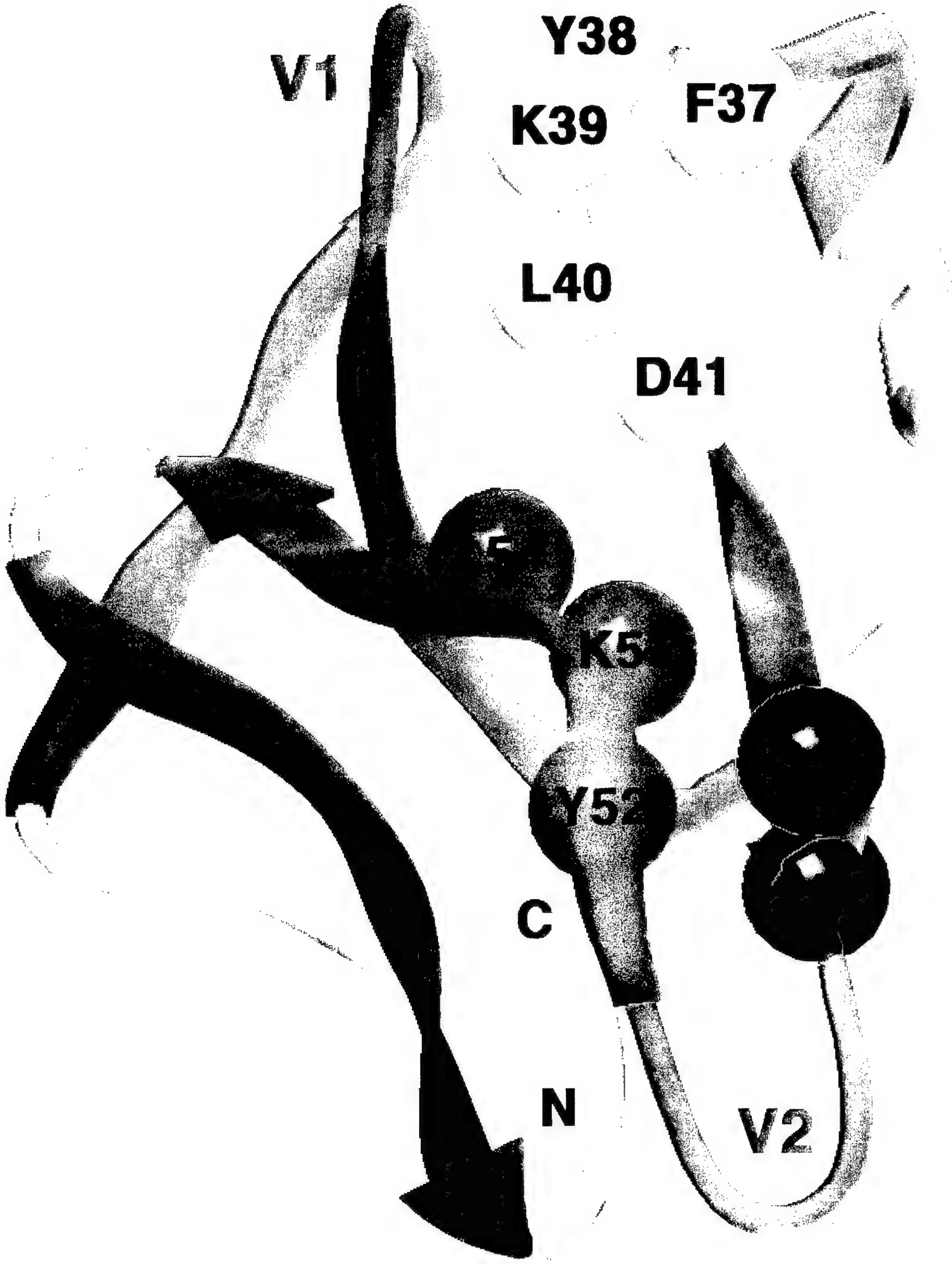


Figure 1. Flow chart of the structure prediction procedures.

Figure 2. Predicted structure for the V1/V2 domain of HIVHXB2R. Residues are numbered from the beginning of the V1/V2 domain, i.e. cysteine 119 is residue 1. The hypervariable V1 and V2 regions are indicated. The V1 loop was shortened for the purpose of the lengthy calculation. Mutations in F37, Y38, K39, L40, D41 and P44, I45 and Y52, K53, L54 reduce or destroy affinity of this domain for HIV-1 neutralizing monoclonal antibodies. Enlarged spheres represent alpha carbons of amino acids involved in maintaining neutralizing epitopes.


V1
Y38
K39
F37
L40
D41
K5
Y52
C
N
V2

The identification of synthetic peptides that block binding of HIV-1 gp120 to CD4 and Fab 12

by

Ruoheng Zhang, Ming-zhu Zhang, Yuan Xie, Edelmira Cabezas and Arnold C. Satterthwait

Department of Molecular Biology, 10666 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

Introduction.

CD4 serves as the primary cellular receptor for HIV-1 [1]. Binding is mediated by a specific interaction with multimeric gp120 on the viral surface[2]. Blocking viral attachment by inhibiting gp120-CD4 binding either by biochemical or immunochemical methods could prevent infection [3]. IgG1b12, a human monoclonal antibody, and tetrameric CD4-IgG (CD4-IgG2), block gp120-CD4 binding and are among the most effective reagents at *in vitro* cross-clade neutralization of primary isolates [4].

IgG12 and CD4 provide important resources for identifying conserved neutralization sites or parts thereof on gp120 that might be mimicked for use as drugs and synthetic vaccines. However, despite considerable attention to CD4/gp120 binding, definition of the CD4 binding site on gp120 has remained elusive [3]. Both reagents bind native protein but not denatured protein [5]. Studies with site-directed mutants of gp120 show that binding is dependent on common, conserved amino acids from different regions of the gp120 primary sequence [6,7]. These observations suggest that IgG1b12 and CD4 bind overlapping, discontinuous epitopes. However, the evidence is indirect and confirmation is required. This could come from X-ray crystallography but no crystal structures have been reported for gp120. Alternatively, one could attempt to establish specific binding to peptide ligands which are needed in any event for the development of synthetic vaccines.

Different claims have been made that gp120 peptides bind CD4 and block CD4/gp120 interaction (8, 9 and 10). Recently, Robey et al. (10) identified a constrained peptide from the C4 region of gp120 that binds CD4 and blocks CD4/gp120 binding. While these results are intriguing, the specificity of this reaction remains to be established. We decided to explore peptide libraries made up from amino aicds believed to comprise the discontinuous epitope on gp120. Since most of the binding affinity of epitopes to antibodies is supplied by several amino acids (1-7, average=3 amino acids), we thought it worthwhile to begin to explore various arrangements of critical amino acids in linear peptides.

Discontinuous epitopes are much more difficult to mimic than continuous epitopes [11]. While discontinuous epitopes and continuous epitopes are composed from similar numbers of critical amino acids, discontinuous epitopes are topologically more complex [12,13]. Peptide libraries, however, are providing powerful new tools for the identification of synthetic leads. Critical amino acids can be randomly positioned in peptide chains and millions of peptides can be screened simultaneously improving chances (11,14,15). In this paper, we report on the results of screening synthetic peptide libraries with Fab 12 and sCD4, expressible binding fragments of IgG1b12 and CD4 respectively. Lead peptides were identified that block gp120-sCD4 binding at about 10 uM (50% inhibition). The same peptides at about 40 uM block gp120-Fab12 binding. These studies indicate a role for acidic and hydrophobic residues in the competitive inhibition of the CD4/gp120 and Fab 12/gp120 interaction.

Experimental

Peptide Synthesis. Individual peptides were synthesized with a ACT350 multiple peptide synthesizer (Advanced ChemTech, USA) using standard Fmoc peptide solid phase peptide synthesis procedures recommended by ACT. Rink amide resin (0.55 mmol/g, ACT) and Fmoc-amino acids (Bachem,CA) were used for these syntheses. Peptides were purified to >95% purity by reverse phase HPLC on C-18 columns using a Gilson HPLC system. The identify of each peptide was confirmed by electrospray mass spectroscopy.

Library design. This was guided by a consideration of general characteristics of discontinuous epitopes and binding data for gp120 (6,7). X-ray crystal structures of

protein-protein interfaces show that 11-19 amino acids from one face contact an opposing face (12). A smaller subset of amino acids, however, termed the functional epitope account for most of the free energy of binding. Functional discontinuous epitopes are made up from an average of three critical amino acids (>20-fold effect on binding, range 1-7 amino acids) and eight contact amino acids (> 2-fold effect on binding, range 4-14 amino acids) within a 20 A diameter. Binding data for Fab12 and CD4 suggest an average level of complexity. Fab12/gp120 HXBc2 binding is nearly abolished by mutations of 358D, 360E and 477D while mutations at 256S/257T, 314G, 384Y, 457D and 470P decrease binding (16). The V1/V2 domain also plays a role since mutations at 177Y, 178K and 179L knockout binding. Surprisingly, however,, deleting V1/V2 from gp120 only reduces binding. CD4/gp120 binding is nearly abolished by mutations of 257T, 358D, 360E, 427W and 457D (6).

Complete peptide libraries with every possible sequence are limited to hexamers by sheer number (14). These libraries are not anticipated to accommodate an average functional discontinuous epitope where C-alpha carbons can be separated by up to 20A. For this reason, we built a shuffled library based on the gp120 amino acids required for binding sCD4 and Fab 12 and added lysines to improve solubility and lengthen the peptides. The following eight building blocks were used: DPE, ST, DY, YW, WP, PG, GK, KD. DPE is a gp120 sequence that includes 358D and 360E. Single amino acids were paired as a device to increase the length of the peptides to at least 12 amino acids. Each of the amino acid blocks was used at six different positions. The length of peptide in the libraries varies between twelve and seventeen amino acids. A dodecamer spans about 40 A in the extended form, double the diameter of an antibody binding site. Thus the peptides can fold back on themselves two or more times to bind a topologically more complex site. Randomly spaced glycines and prolines facilitate peptide turns that could aid in binding.

Library Synthesis. Peptide libraries were synthesized with the ACT350 multiple peptide synthesizer using the divide, couple and recombine method (DCR method) (17,18). The ACT350 performs up to 96 independent coupling reactions in individual reaction tubes per coupling cycle. NovaSyn TGR resin (0.20 mmol/g, Novabiochem) was used as the solid support for these syntheses. In the DCR method, one or more amino acids are coupled to 20 mg resin/synthesis tube using standard Fmoc chemistry. The coupling of each amino acid is driven to completion by an excess of amino acid and spot checked for

completion with the ninhydrin assay. After one or a block of amino acids are coupled, resins with different sequences from different tubes are combined and redistributed for subsequent rounds of amino acid additions. Upon completion of the synthesis, peptide libraries are cleaved from individual tubes with a trifluoroacetic acid cocktail (75mg phenol, 50uL thioanisole, 50 uL H_2O and 25 uL ethanedithiol in 1 mL TFA), precipitated, washed twice with diethyl ether and lyophilized. The peptides were reconstituted in HBS buffer : (10mM HEPES (PIERCE), 150 mM NaCl (Fisher), 3.4 mM EDTA (Sigma), 0.005%(v/v) surfactant P20 (Pharmacia Biosenser AB) in distilled H_2O). The pH was adjusted to 7.4 with 2.5 N NaOH(Fisher) to yield a final peptide concentration of 2.5 mg/ ml (1.6mM based on an average molecular weight) for binding studies.

CD4 competition ELISA. In this assay, peptide blocks horse radish peroxidase labeled (HRP) gp120 IIIB from binding to human sCD4 (D1D2 domains, MW = 45,000, SmithKline and Beecham) adsorbed to a plastic surface. Both sCD4 and HRP-gp120 were titrated under assay conditions to ensure that minimum quantities were used in the assay. Operations were carried out at room temperature unless otherwise stated.

sCD4 was adhered to the wells of medium binding plates (Costar) by adding 50 ng sCD4 in 50 ul PBS buffer per well (Costar) and incubating overnight at 4 C. The wells were washed two times with 0.05% Tween/phosphate buffer saline (PBS), blocked with 200 ul 1% bovine serum albumin (BSA)/0.05%Tween/PBS for 1 hr and washed again. Peptide (12-1600 uM) in 45 ul 1% BSA/HBS was added to each well and preincubated for 1 hr. Then 50 ng HRP-labeled baculovirus gp120 IIIB (Intracel) in 5 ul 1% BSA/HBS buffer was added to each well and incubated for 1 hr at 37 C. The reaction wells were washed 5 times with 0.05% tween/PBS and 100ul 3,3',5,5'-tetramethyl benzidine solution (PIERCE) added for 30 min. The reaction was stopped with 100ul of 2N H_2SO_4 per well and absorption measured at 450 nm after an appropriate time using a Spectramax 250 ELISA plate reader.

Fab12 competition ELISA. In this assay, peptide is blocks Fab 12 from binding to gp120 IIIB (Intracel). Gp120 IIIB is initially bound to D7324 (International Enzymes) adsorbed to a microtiter well. Fab 12 was supplied by Dr. Carlos Barbas. D7324 is purified polyclonal sheep antibody that binds the C-terminus of gp120. Operations were carried out at room temperature unless otherwise stated.

First, 1 ug D7342/ 50 uL 0.1M $NaHCO_3$, pH 9.6 per well was coated on a medium binding 96 well microtiter plate (Costar) overnight at 4 C. The plate was washed twice with deionized water, blocked with 5% blotto (5% non-fat milk powder, 0.02% antifoam A in PBS) for 1 hr and washed. Then 50ng gp120 IIIB (Intracel) in 100 ul 5% blotto was added to each well for 1 hr and the plate washed two times as above. Next 45 ul peptide in HBS buffer was preincubated with Fab12 (5ng/5 ul) in 1% BSA/PBS for 1 hr and added to each well. After incubation at 37 C for 1 hr, the plate was washed five times as above and 100ul alkaline phosphatase labeled goat antihuman IgG(Fab')$_2$ (PIERCE) was added to the wells and incubated at 37 C for 1 hr. After washing the plate five times, 100 ul developing solution (5 mg p-nitrophenyl phosphate in 5 ml developing buffer made from 50 ml diethanolamine (Sigma), 50 mg $MgCl_2$, 97.5 mg NaN_3) was added and absorbance measured at 415nm after an appropriate time with a Bio-tek EL309 microplate reader.

Surface plasmon resonance measurements. Real time binding of peptides or gp120 to sCD4 or control BSA was measured with a BIAcore 1000 on research grade CM5 sensor chips (Pharmacia BiosensorAB). The amino coupling reagent, N-hydroxyl-succinimide (NHS), N-ethyl-N'-(3-diethyl aminopropyl) (EDC) were from PIERCE and ethanolamine hydrochloride was from Sigma. HBS buffer used to dilute and run samples was prepared as described above. Reactions were carried out at 25 C.

Human sCD4 (D1D2 domains, Intracel or SmithKline and Beecham) and gp120 IIIB(Intracel) were immobilized on individual sensor surfaces according to the manufacturer at a flow rate of 5 ul/min. (Fig. 1). First, the instrument was equilibrated with HBS buffer and the sensor chip surface activated (a) with 30ul of an equal volume of 0.1 M NHS and 0.1M EDC. Then the ligand (30ul, 30 ug protein/ml 10 mM sodium acetate) was injected (b). Next the unreacted sites were capped (c) with 30ul, 1M ethanolamine in water, pH8.5. Finally, noncovalently bound ligand was washed from the surface (d) by injecting 15 ul 0.1M H_3PO_4.

Fifty ul of peptide (1.6 mM) in HBS buffer was injected over sCD4 immobilized on the sensor chip surface at a flow rate of 10 ul/min and the increase in response units (RU) recorded. For inhibition experiments, sCD4 receptor was incubated with peptide for 30 min prior to injection over gp120 immobilized on the sensor chip surface. CD4 and gp120 bound to the sensor surface were freed from ligand by regenerating with 15 ul 0.1 M H_3PO_4.

Results.

First library. Initially, a soluble combinatorial library of 229,376 peptides in 56 pools were synthesized and screened. In this library, each of eight building blocks (DPE, ST, DY, YW, WP, PG, GK, KD) were initially added in eight separate reactions to equal amounts of resin that were mixed to give X_6-resin. This mixture was divided into eight equal parts and the process repeated to give X_5X_6-resin and eventually $X_3X_4X_5X_6$-resin consisting of equimolar quantities of 4,096 peptides. The last two positions (O_1,O_2) were filled in separate reactions to yield 56 pools of Ac-$O_1O_2X_3X_4X_5X_6$-NH_2 where the first two blocks were defined in the individual pools. The peptide pools were individually cleaved from the resin and dissolved in assay buffer (2.5 mg/ml). Each of the 56 pools contained 4, 096 peptides. Each peptide in each pool is 1,600 uM. Each peptide pool was screened for binding to sCD4 using surface plasmon resonance measurements (SPR) and competition ELISA.

Peptide mixtures rich in DPE and hydrophobic residues (YW,WP and PG) showed weak affinity for sCD4 using SPR measurements (Table 1, Fig. 2). The increase in response units (RU), a measure of surface plasmon resonance, is specific for the DPE peptides and proportional to immobilized sCD4 ruling out the bulk effect. (Fig. 3). The preference was for DPE at either the O1 or O2 positions with YW, WP and PG occupying the adjacent slot.

The same DPE peptides that bind weakly to sCD4 also blocked binding of labeled gp120 to sCD4 and partially inhibited binding of Fab12 to gp120 in competition ELISAs (Table 2). We confirmed this result using SPR. Peptide libraries, acetyl-YWDPEXXXX-NH_2, acetyl-PGDPEXXX-NH_2 and acetyl-PGKDXXXX-NH_2 were each mixed in individual reactions with sCD4 and injected over gp120 immobilized on the sensor surface. While sCD4 bound gp120, acetyl-YWDPEXXXX-NH_2 blocked binding (Fig. 4). Acetyl-PGDPEXXX-NH_2 blocked binding in the same way but acetyl-PGKDXXXX-NH_2 did not (not shown). This confirms that the DPE peptides block sCD4/gp120 binding.

Second library. Defining the fourth block. Data from the first library indicated that Ac-X_1DPEX_3XXX-NH_2, (X_1 and X_3 =YW, WP,PG) block sCD4/gp120 binding best. To define the fourth position and confirm the first and third positions, four libraries, acetyl-

YWDPEYWO4XX-NH_2, acetyl-YWDPEWPO4XX-NH_2, and acetyl-YWDPEPGO4XX-NH_2 and acetyl-PGDPEYWO4XX-NH_2 were synthesized and tested for binding to sCD4 using SPR (Fig. 5) and inhibition of the gp120/CD4 reaction using competition ELISA (Fig. 6). The peptide mixtures (2.5 mg/ml) were soluble but close to saturation. SPR measurements were made at 2.5 mg/ml while competition ELISA was carried out with 0.6 mg/ml. Relative concentrations of the peptide mixtures were checked with HPLC confirming equimolar quantities and solubility of the peptides. Both tests indicated that acetyl-YWDPEYWYWXX-NH_2. bound sCD4 best and blocked sCD4/gp120 best..

Our strongest case for specific binding of the peptide libraries to sCD4 relied on the initial observation that DPE peptides were preferred. Mutations of $D^{368}PE^{370}$ abrogate binding of gp120 to sCD4 and Fab 12. To study how DPE in the peptide libraries contributes to the binding of sCD4, two peptide mixtures, acetyl-YW**RPE**YWYWXX-NH_2 and acetyl-YW**AAA**YWYWXX-NH_2, were synthesized for comparison with acetyl-YW**DPE**YWYWXX-NH_2. D was replaced with R in the first library since D^{368}/R mutation of gp120 IIIB knocks out binding to sCD4 (6). DPE was replaced with AAA in the second peptide. Both peptides suffered from solubility problems. Nonetheless, SPR measurements showed that lower concentrations of acetyl-YWAAAYWYWXX-NH_2 bound sCD4 but then dissociated relatively slowly (Fig.7). Unexpectedly, it appeared that acetyl-YWAAAYWYWXX-NH_2 and acetyl-YWRPEYWYWXX-NH_2 increased binding of gp120 IIIB to CD4 (Table 3). While this might imply that the peptides act as a non-specific glue to bind sCD4 and gp120 together, we suspect that this increase arose because the peptide precipitated on the plate and adsorbed additional amounts of gp120 non-specifically.

Third library. Defining the fifth block. The peptide library based on, acetyl-YWDPEYWYWO5X-NH_2 (O5= DPE, ST, DY, YW, WP, PG, GK or KD) was synthesized. SPR measurements (Fig. 8) indicated that Ac-YWDPEYWYWYWX-NH_2, and Ac-YWDPEYWYWWPX-NH_2 bound sCD4 better than the other peptides. Competition ELISA suggested that Ac-YWDPEYWYWWPX-NH_2 blocked CD4/gp120 binding best (Fig. 9). However, Ac-YWDPEYWYWYWX-NH_2 may have undergone some precipitation that reduced its effective concentration and apparent competition.

Fourth library. Defining the sixth block of library. Eight individual peptides, acetyl-YWDPEYWYWWPO$_6$-NH$_2$ (O$_6$=DPE, ST, DY, YW, WP, PG, GK, KD) were synthesized to define the last block. Each peptide was purifed to >95 % purity by high pressure chromatography and their identitities confirmed by mass spectroscopy. All of the peptides were soluble at 1.25 mg/ml except for the peptide where O$_6$ = YW. Concentrations were confirmed by HPLC. Both SPR measurements (Fig. 10) and competition ELISAs (Fig. 11) were carried out with sCD4. It is clear from these experiments that acetyl-YWDPEYWYWWPWP-NH$_2$ binds sCD4 best. It also blocks sCD4/gp120 best (50% blocking at about 10 uM). The low apparent affinity of peptide-O$_6$ = YW for sCD4 and less effective blocking is due in the first instance to a lower concentration. The same peptides were tested in competition ELISA for blocking Fab12/gp120 binding(Fig. 12). Again, it was found that acetyl-YWDPEYWYWWPWP-NH$_2$ blocked best (about 50% blocking at 38 uM). We refer to acetyl-YWDPEYWYWWPWP-NH$_2$ in subsequent work as Peptide 1.

While an IC$_{50}$ = 10uM for the sCD4/gp120 interaction would generally be considered significant, we were concerned by the preponderance of hydrophobic amino acid residues in Peptide 1. In addition, SPR measurements were high (Fig 10) suggesting that more than one peptide might bind or stick to sCD4 albeit transiently. We are and were concerned about the possibility of non-specific effects. For example, the peptide might bind and destroy sCD4 or gp120 IIIB. To test for this, we pretreated plate bound sCD4 with Peptide 1, washed the peptide from the wells and added labeled gp120 IIIB. The gp120 bound pretreated sCD4 to the same extent that it bound untreated gp120 indicating that Peptide 1 has no effect on the structure of sCD4. We also showed that pretreatment of D7324 bound gp120 IIIB with Peptide 1 had no effect on its binding of Fab 12. In another control, we showed that Peptide 1 did not block MAb 59.1, a V3-loop antibody, from binding gp120, at concentrations that it completely blocked binding of sCD4.

DPE dependent binding ? In a test for specificity, we shuffled the DPE amino acids within the confines of the second block of Peptide 1. The C-terminal O$_6$ block was deleted for this test. Three individual peptides, acetyl;-YW**DPE**YWYWWP-NH$_2$, acetyl-YW**PDE**YWYWWP-NH$_2$ and acetyl-YW**EDP**YWYWWP-NH$_2$ were synthesized and purified. Each of the peptides blocked gp120/Fab 12 binding at similar concentrations (Fig. 13). A larger set of peptides was prepared where D, P and E was repositioned within Peptide 1. These peptides blocked the sCD4/gp120 interaction with IC$_{50}$ = 10 - 100 uM

(Fig. 14). While the range of inhibitory constants is intriguing, a slight precipitate was noted in one day old stock solutions for Peptide 1 and acetyl;-YW**DPE**YWYWWPYW-NH_2 suggesting that the concentrations of these peptides might be lower than indicated. This places an upper limit on the recorded IC_{50} values and may account for the scatter of values. Despite these uncertainties, it seems clear that binding does not require sequential DPE. It remains to be seen whether a positional preference for D, P and E can be established.

Discussion

We decided to explore peptide libraries made up from amino acids believed to comprise the discontinuous epitope on gp120. Since most of the binding affinity of epitopes to antibodies is supplied by several amino acids (1-7, average = 3 amino acids), we thought it worthwhile to begin to explore various arrangements of critical amino acids in linear peptides. While different library designs can be considered, we chose one based on the characteristics of discontinuous epitopes and amino acids from gp120 believed to make up the Fab 12 and CD4 binding site(s). Although more needs to be done, several features are beginning to emerge. Hydrophobic peptides rich in tryptophans that are solublized by aspartic and glutamic acid residues block sCD4/gp120 binding with IC_{50} = 10uM.

Peptide 1 binds sCD4 albeit transiently and blocks the sCD4/gp120 interatction. It is reasonable but not proved that peptide 1 binds the gp120 binding site on sCD4. Mutagenesis studies carried out on sCD4 have implicated a patch of exposed residues on the D1 domain of sCD4 (**Lys 29**, **Lys 35**, Phe 43, Leu 44, **Lys 46**,Gly 47 and **Arg 59**) that could bind gp120 (19). The positively charged residues (in bold) on sCD4 could align themselves with the negatively charged residues on gp120 (**Asp368**, **Glu370** and **Asp457** and possibly **Asp477**) required for binding CD4. **Asp368**, **Glu370**, **Asp477 and possibly Asp457** are required for binding Fab 12 (6). Examination of the 3D-structure for CD4 (20) shows how the amino acid sequence folds back on itself to bring the β-carbons of Lys 29 and Lys 35 within 5 A of each other. This is close enough for the positively charged side chains of Lys 29 and Lys 35 on sCD4 to be in position to bind negatively charged side chains on D368 and E370 on gp120. It is difficult to ignore the juxtaposition

of positive charges on CD4 and negative charges on gp120. The peptides emerging from our libraries may reflect this possibility.

Recently, Robey *et al.* (18) have reported that a peptomer, an oligomerized form of the C4 peptide KIKQIINMWQEVGKAMYA-NH_2, binds sCD4 very effectively when it is adsorbed to a titer well. They also report that sCD4 binds a peptomer based on the comparable HIV-2 sequence, HIEQIINTWHKVGKNVYL-NH_2 and that the HIV-1 monomer, KIKQIINMWQEVGKAMYA-NH_2, in 0.035% Brij blocks the CD4/gp120 interaction with IC_{50} = 42 uM. Brij is required for blocking. Since binding correlated with the observation of alpha helix by circular dichroism and the two peptides share hydrophobic residues, they proposed that sCD4 binds the shared hydrophobic face of the predicted amphipathic alpha helix (shared hydrophobic residues are in bold). They also report that alpha helix formation and blocking of the sCD4/gp120 reaction is dependent on an intact tryptophan.

How do the results of Robey *et al.* (18) compare with our results ? First, Peptide 1 and the C4 monomer in Brij both block sCD4/gp120 binding at similar concentrations. Robey *et al,* proposed site of interaction on the peptomer is a hydrophobic surface and involves a tyrptophan. Peptide 1 is also hydrophobic and rich in tryptophans. Peptide 1 carried two acidic residues while the peptomers each carry one. However, the sequences of Peptide 1 and the C4 monomer differ considerably. Although the situation remains unclear how either the peptomer or Peptide 1 block sCD4/gp120, these peptides raise provocative questions which should aid in continued research.

References

1. Dalgleish, A.G., Beverely, P.C.L., Clapham, P.R., Crawford, D.G., Greaves, M.F., Weiss, R.A., (1984) Nature(London) 312, 763~766.

2. Klatzmann, D., Champagne, E., Charmaret, S., Gruest, J., Guetard, D., Hercent T., Gluckman, J. -C., Montagnier, L., (1984) Nature (London) 312,767~768.

3. Moore, J.P. and Sweet, R.W., (1993). The HIV gp120-CD4 interaction: A target for pharmacological or immunological intervention?, Perspective in Drug Discovery and Design 1 235-250 ESCOM

4. Trkola, A. et al., (1995). Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG, J. Virol. 69: 6609-6617 .

5. Ditzel, H. J. et al., (1995). Neutralizing recombinant human antibodies to a conformational V2- and CD4-binding site-sensitive epitope of HIV-1 gp120 isolated by using an epitope-masking procedure, J. Imm. 154: 893-906

6. Olshevsky, U. et al., (1990) Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4 receptor binding, J. Virol 64:5701

7. Roben, P., Moore J.P., Thali, M., Sodroski, J., Carlos, B., Burton, D., (1994) Recognition properties of a panel of human recombinant Fab fragments to the CD4 binding site of gp120 showing differing ability to neutralise human immunodeficiency virus type 1, J. Vir., 68(8) 4821~4828.

8. Tamamura, H. et al.,(1996). Interaction of an anti-HIV peptide, T22, with gp120 and CD4, BBRC 219, 555

9. Ramsdale, T. E. et al. (1993). Verification of the interaction between peptide T and CD4 using surface plasmon resonance, FEBS 333: 217

10. Robey, F. A. et al. (1996). A synthetic conformatonal epitope from the C4 domain of HIV gp120 that binds CD4, J. Biol. Chem. 271: 17,990

11. L.L.C. Bonnycastle et al., (1996). Probing the basis of antibody reactivity with a panel of constrained peptide libraries displayed by filamentous phate, J. Mol. Biol. 258:747

12. Janin, J. and Chothia, C., The structure of protein-protein recognition sites, J. Biol. Chem., 265:16027-16030 .

13. Lei, J. Fendly, B. M. and Wells, J. A., (1990). High resolution functional analysis of antibody-antigen interactions, J. Mol. Biol. 226:851 (1992).

14. Houghten, R.A., Pinilla, C., Blondelle, S.E., Apple, .R., Dooley, C., Cuervo, J.H., (1994) Nature,354(7) 84~86.

15. Lam K.S., Salman S.G., Hersh E.M., Hruby, V. J., Kazmierski, W. M., Knarp, R.J., (1991) Nature, 354(7) 82~84.

16. Updated list of gp120 amino acids required for binding Fab12 provided by Dr. Carlos Barbas. See also ref. 7.

17. Furka, A. et al. (1991). Int. J. Pept. Protein Res. 37: 487-493

18. Pinilla, C. et al., (1995). A review of the utility of soluble peptide combinatorial libraries, Biopolymers 37: 221-240

19. Ryu, S-E. et al. (1994) Structure of an HIV and MHC binding fragment from human CD4 as refined in two crystal lattices, Cur. Biol. 2: 59

20. Coordinates for the D1/D2 domain of sCD4 solved by Wang, J. et al & Harrison, S.C. (1990) Crystal structure of an HIV-binding recombinant fragment of human CD4. Nature 348: 419 have been deposited in the Brookhaven Database in the file: pdb3cd4.ent.

$O_1 \backslash O_2$	DPE	ST	DY	YW	WP	PG	GK	KD
DPE		558	233	**790**	**860**	**918**	554	403
ST	258		98	147	141	410	353	110
DY	339	211		635	202	445	239	67
YW	**411**	282	468		132	99	508	252
WP	**532**	117	165	132		343	361	131
PG	**718**	91	211	84	148		202	190
GK	293	122	161	248	152	171		103
KD	47	74	52	182	360	92	251	

Table 1. Surface plasmon resonance measurements for the binding of peptide library 1, acetyl-O_1O_2XXXX-NH_2 (2.5 mg/mL), to sCD4 immobilized on the sensor chip. Response units (RU) were measured at 5 seconds after the injection of peptide was completed and recorded in the above boxes. For example, RU 790 was observed for DPEYWXXXX-NH_2 as shown in Fig. 1 and recorded in the appropriate box above.

O_1O_2	BIAcore/CD4 (RU)	ELISA/CD4 (% Bound)	ELISA/Fab12 (% Bound)
DPEWP	869	0	47
DPEYW	790	5	68
YWDPE	411	1	54
WPDPE	532	2	60
PGDPE	718	63	65
PGKD	190	>100	68
KDDY	52	>100	94

Table 2. A comparison of SPR measurements for ,Ac-O_1O_2XXXX-NH_2 binding to immobilized sCD4 (RU) and competitive inhibition of CD4/gp120 and Fab12/gp120 interactions. Peptide libraries were 2.5 mg/ml.

Concentration (ug/ml) Peptide	1250	600	300	150	75
Ac-YWDPEYWYWXX-NH_2	6	28	52	ND	ND
Ac-YWRPEYWYWXX-NH_2	ND	235	165	122	118
Ac-YWAAAYWYWXX-NH_2	ND	404	449	452	379

Table 3. Competition ELISA of peptide mixtures with gp120 IIIB for sCD4. % gp120 bound to sCD4 on plate in the presence of peptide relative to a control without peptide.

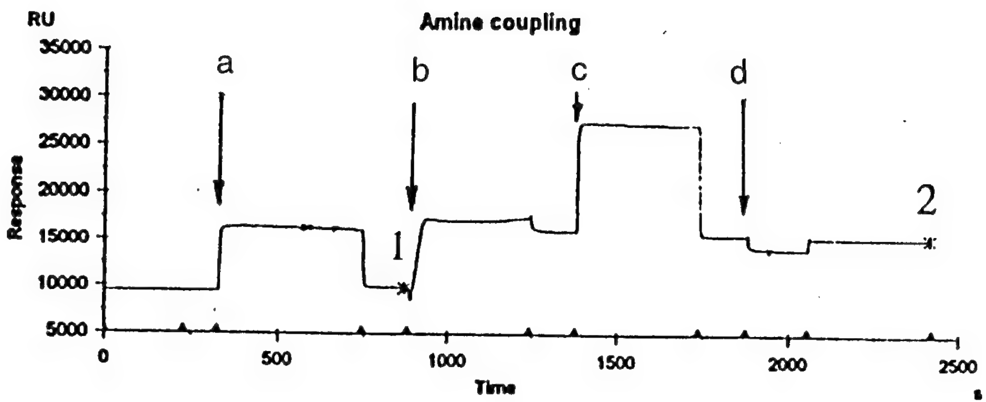


Figure 1. Sensorgram recording the immobilization of sCD4 on a CM5 sensor chip. The difference in response units between point 1 and point 2 (5,523 RU) provides a measure of immobilized sCD4.

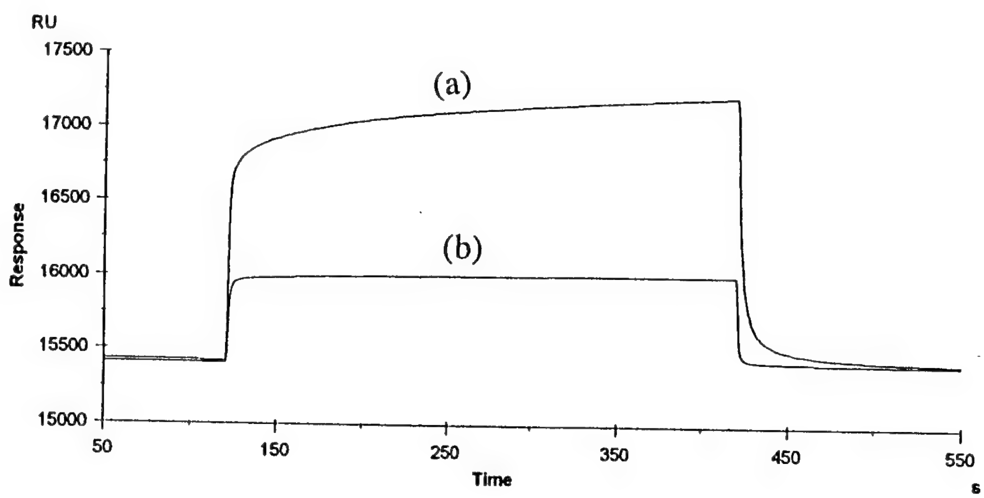


Figure 2. Sensorgram for binding peptide mixtures (50 uL, 2.5 mg/ml) of (a) acetyl-DPEYWXXXX-NH_2 and (b) acetyly-PGWPXXXX-NH_2 to 5,523 RU of immobilized of sCD4. Flow rate = 10 uL/min.

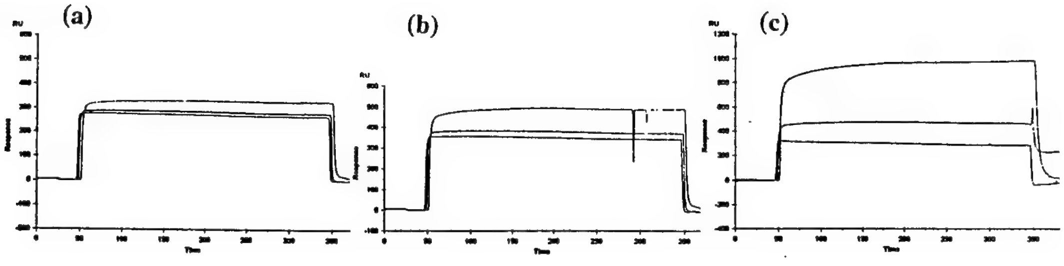


Figure 3. Sensorgrams for (a) acetyl-KDDYXXXX-NH_2, (b) acetyl-PGDPEXXXX-NH_2, and acetyl-YWDPEXXXX-NH_2 (50 uL, 2.5 mg/ml) passed over different quantities of immobilized sCD4 (25, 1071 and 3987 RU). Flow rate = 5 uL/min.

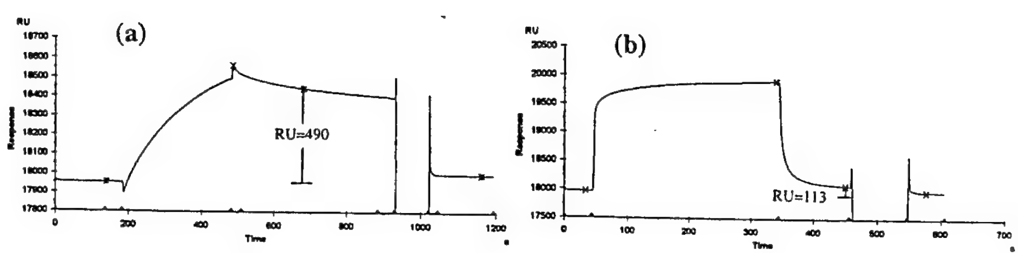


Figure 4. Sensorgram showing binding of sCD4 (30 uL, 12.5 ug/mL) to immobilized gp120 IIIB (a) without and (b) with preincubation with acetyl-YWDPEXXXX-NH_2 (2.5 mg/ml) for 1 h at rt.

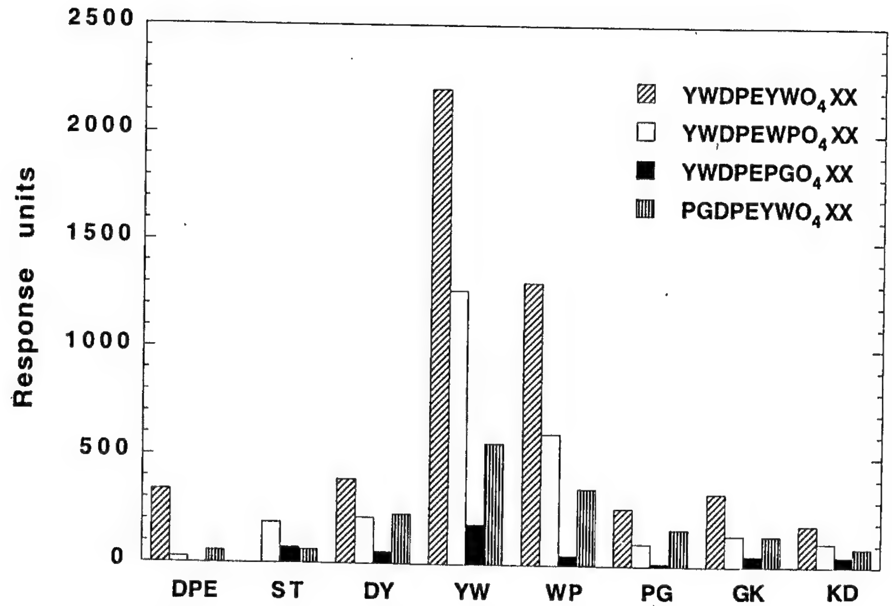


Figure 5. Response units from binding of Ac-YWDPEYWO4XX-NH2 , Ac-YWDPEWPO4XX-NH2 , Ac-YWDPEPGO4XX-NH2 and Ac-PGDPEYWO4XX-NH2 to immobilized sCD4. RUs are measured five seconds after the peptide injection has been completed. O4 = DPE, ST, DY, YW, WP, PG, GK and KD.

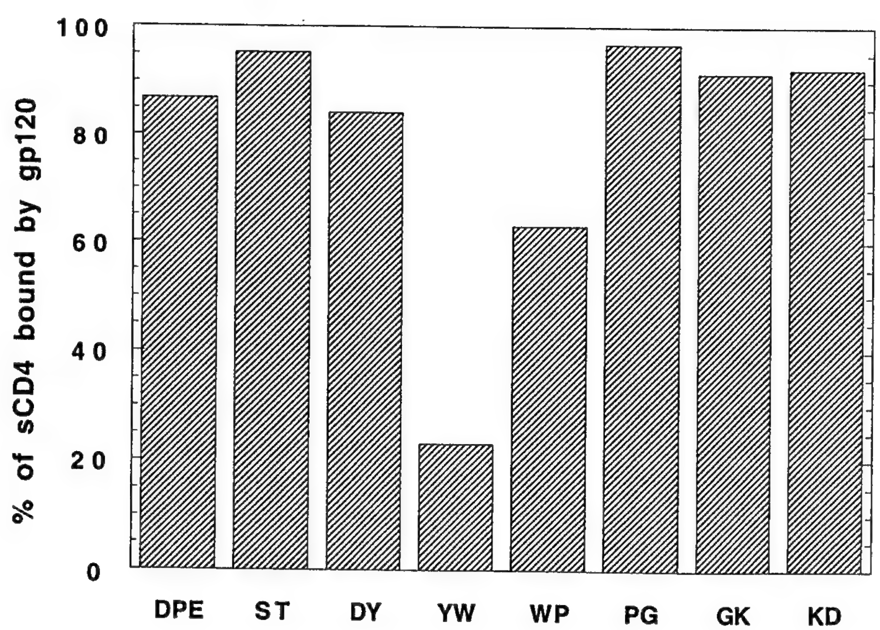


Figure 6. Competition ELISA with Ac-YWDPEYWO4XX-NH2 (0.6 mg/ml) for HRP-labeled gp120 IIIB binding to sCD4. O4 = DPE, ST, DY, YW, WP, PG, GK and KD.

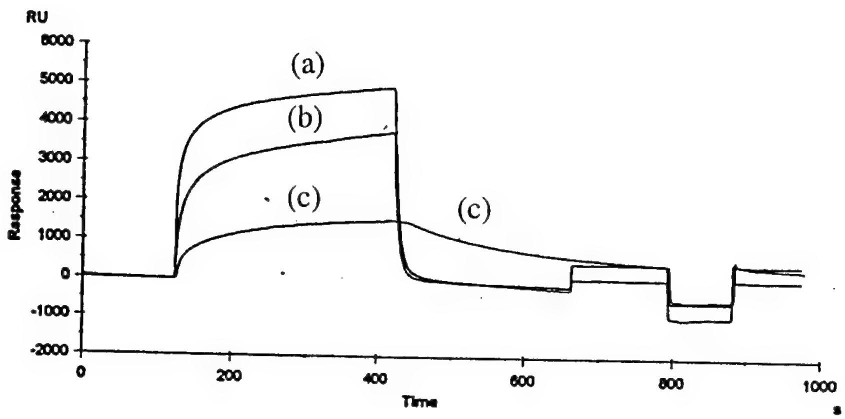


Figure 7. Sensorgram for binding (a) acetyl-YWDPEYWYWXX-NH_2, (b) acetyl-YWRPEYWYWXX-NH_2 and (c) acetyl-YWAAAYWYWXX-NH_2 to immobilized sCD4. The latter peptide sticks more tightly to sCD4 than other peptides.

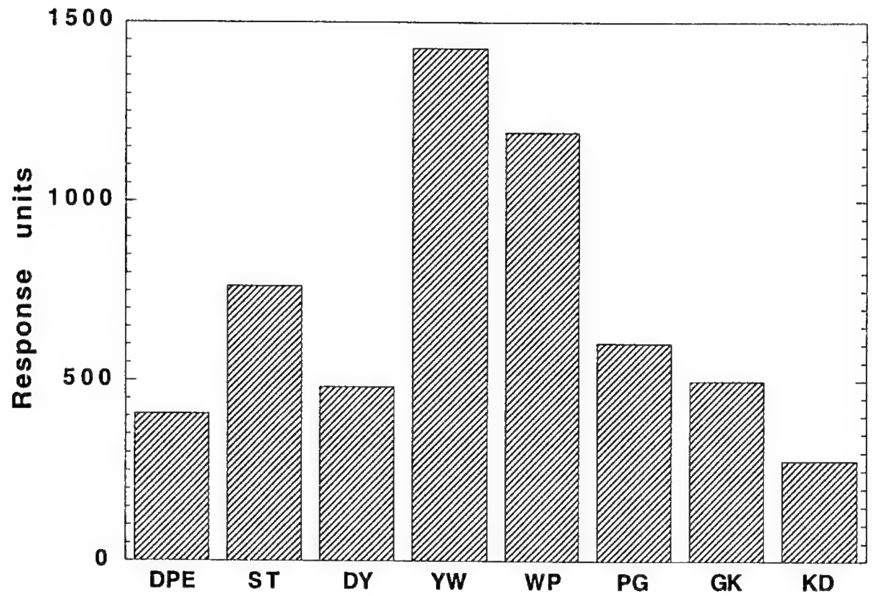


Figure 8. Response units from binding acetyl-YWDPEYWYWO5X-NH_2 (2.5 mg/ml) to immobilized sCD4. RUs are measured five seconds after the peptide injection has been completed. O_5 = DPE, ST, DY, YW, WP, PG, GK and KD.

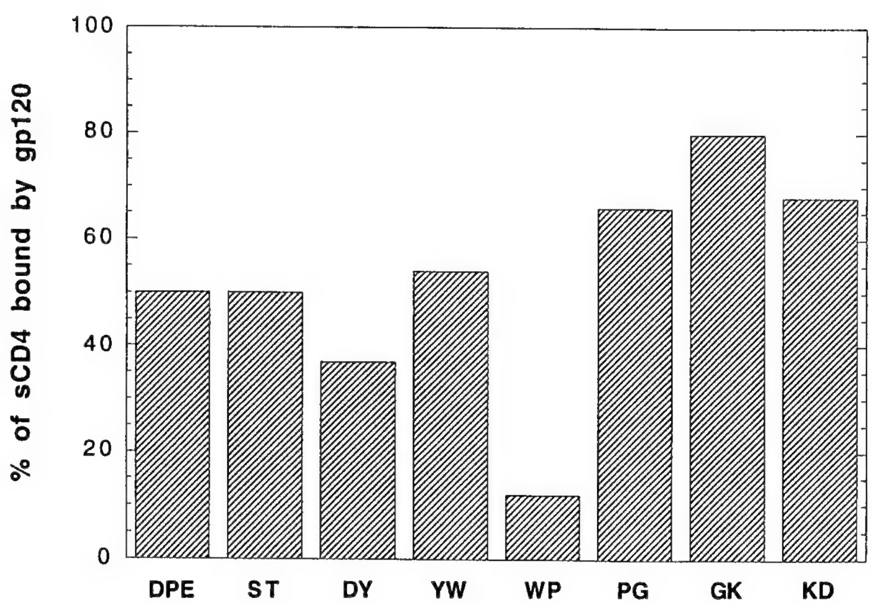


Figure 9. Competition ELISA with Ac-YWDPEYWYWO5X-NH_2 (0.6 mg/ml) for HRP-labeled gp120 IIIB binding to sCD4. O_5 = DPE, ST, DY, YW, WP, PG, GK and KD.

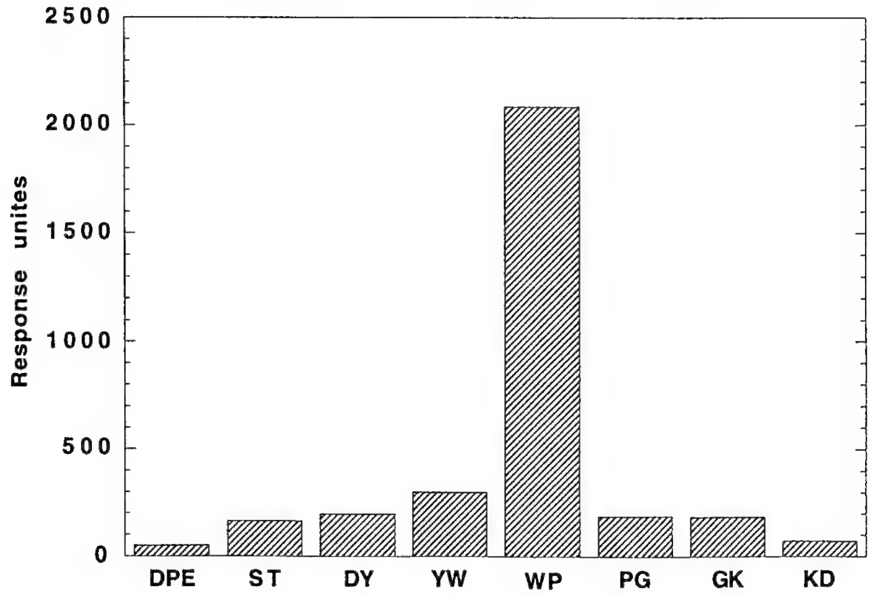


Figure 10. Response units from binding acetyl-YWDPEYWYWWPO6-NH2 (1.25 mg/ml) to immobilized sCD4. RUs are measured five seconds after the peptide injection has been completed. O6 = DPE, ST, DY, YW, WP, PG, GK and KD. Reduced RU for YW is due to a lower concentration for this peptide.

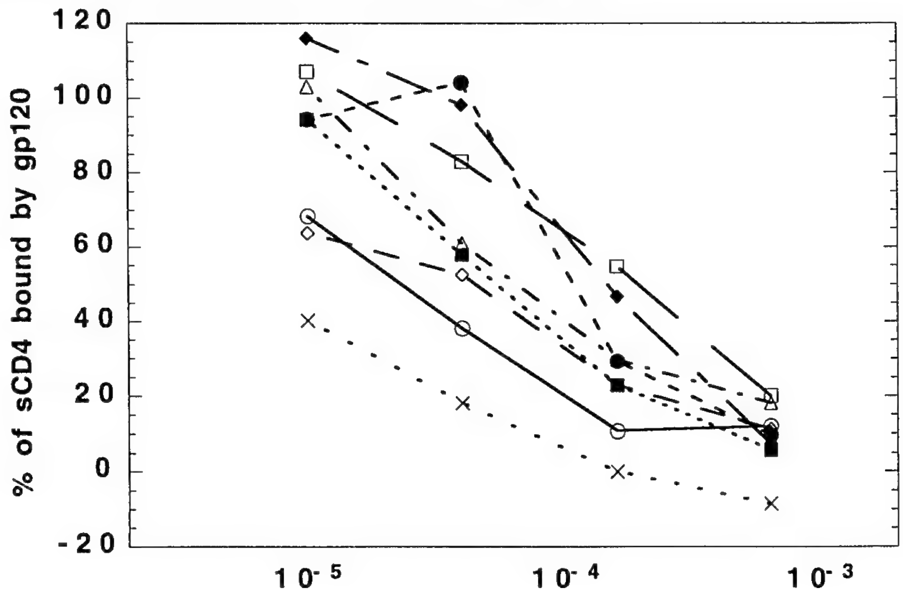


Figure 11. Competition ELISA with Ac-YWDPEYWYWWPO6-NH2 for HRP-labeled gp120 IIIB binding to sCD4. O6 = —o— DPE —□— ST —◇— DY - -●- - YW

- -×- - WP —◆ - PG — △— GK - - -■- - - KD

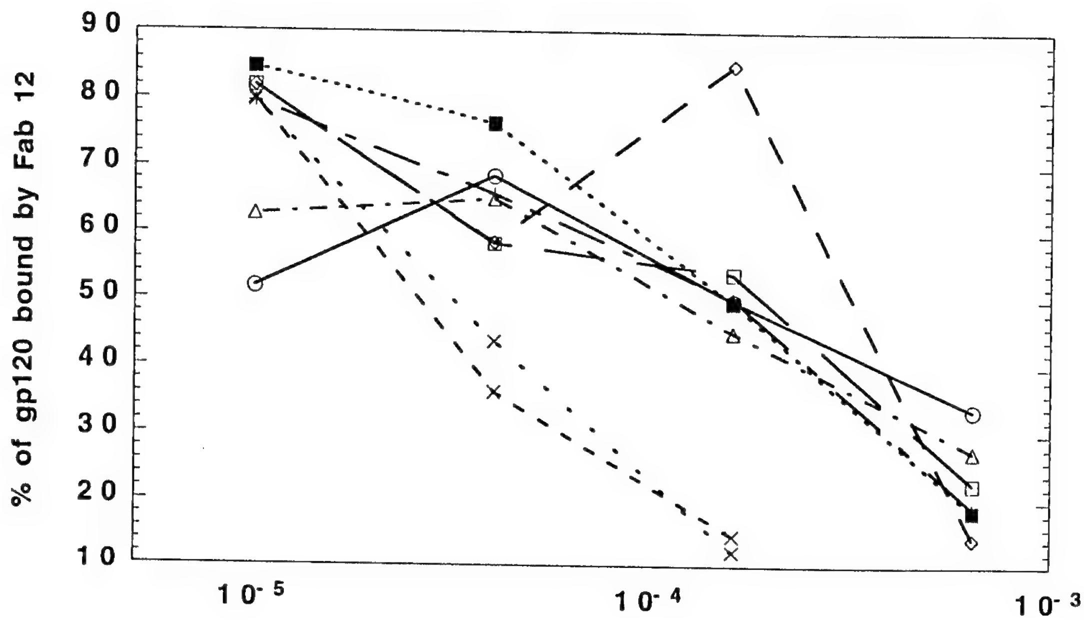


Figure 12. Competition ELISA with Ac-YWDPEYWYWWPO$_6$-NH$_2$ for Fab12 binding to gp120 MN captured on D7324. O$_6$ = —⊖— DPE —⊟— ST —◇— DY - -×- - YW · -×- · WP —+ - PG – △– GK ···■··· KD

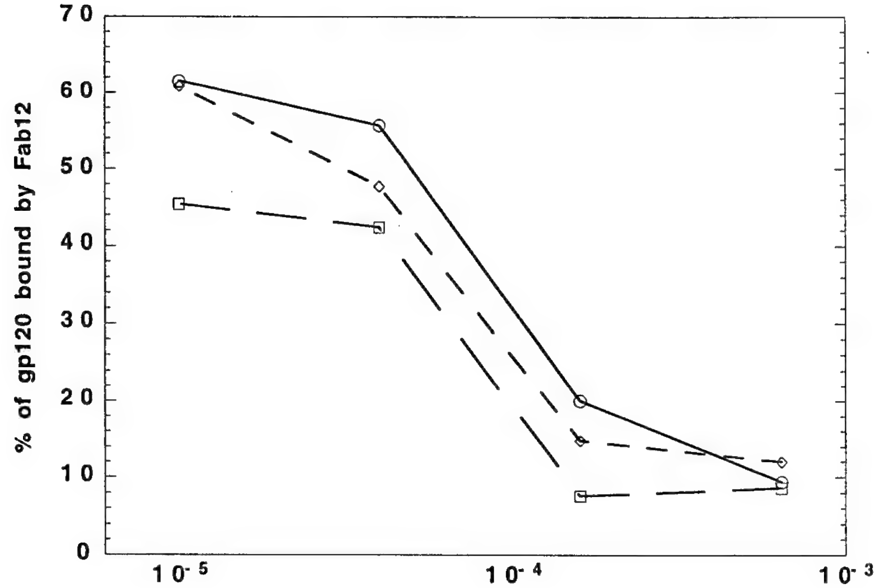


Figure 13. Competition ELISA with —○— Ac-YWDPEYWYWWP-NH2, —◇— Ac-YWDPEYWYWWP-NH2 and —□— Ac-YWDPEYWYWWP-NH2 for Fab12 binding to gp120 IIIB captured on D7324.

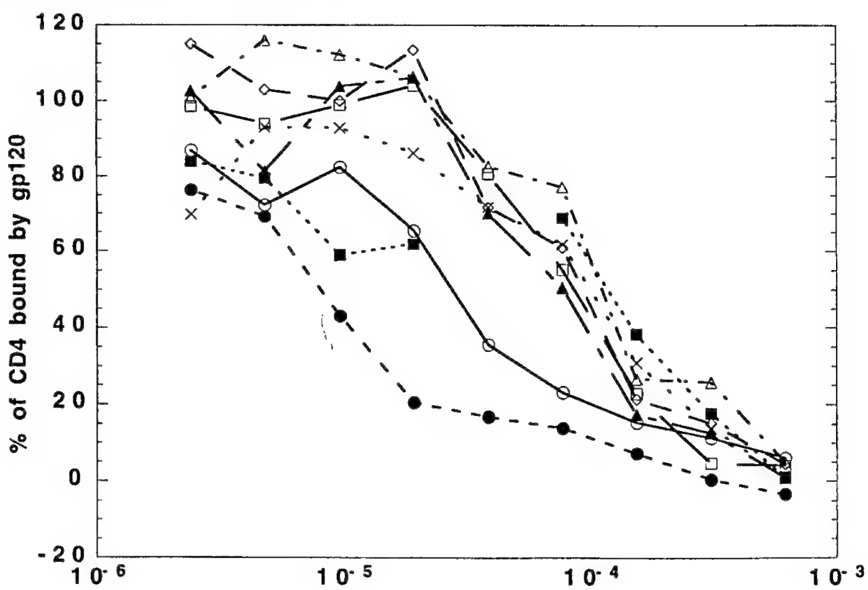


Figure 14. Competition ELISA with peptides for sCD4 binding to gp120 IIIB.

—○— Ac-YW**DPE**YWYWWPYW-NH2, ·×· Ac-YW**PE**YWYW**D**WPWP-NH2,
—□— Ac-YW**DPE**YWYWWPWP-NH2, —▲— Ac-YW**EP**YWYW**D**WPWP-NH2,
—◇— Ac-YW**DP**YWYW**E**WPWP-NH2, –△– Ac-YW**D**YW**P**YW**E**WPWP-NH2,
-●- Ac-YW**PD**YWYW**E**WPWP-NH2, ···■··· Ac-YW**EAP**YWYW**D**WPWP-NH2,

The Immunogenicity of a Constrained Peptide from the Third Constant Region of HIV-1 gp120

by

Arnold C. Satterthwait, Shao-qing Chen and Edelmira Cabezas

Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

An important goal in HIV vaccine research is the identification of neutralizing epitopes that can serve as a basis for a synthetic peptide vaccine. This has led researchers to focus on peptides from the "tip" of the third variable region (V3) of the envelope glycoprotein (gp120) on HIV-1 since they stimulate potent neutralizing antibodies. However, rapid mutation of the virus provides escape routes that limit the potential effectiveness of V3 peptides (1).

On the other hand, IgG1b12, a human monoclonal antibody isolated from a phage combinatorial antibody library, and tetrameric CD4-IgG (CD4-IgG2) are considerably more effective at cross-clade neutralization of HIV-1 primary isolates (2). Both reagents block gp120 from binding to the cellular receptor CD4 (3). Cross-clade neutralization suggests that the CD4 binding site on gp120 is conserved making it a preferred target. However, IgG1b12 and CD4-IgG2 bind native gp120 rather than denatured protein indicating strict conformational requirements (4,5). Epitope mapping with gp120 site-specific mutants further suggest that the expressible binding fragments, Fab12 and sCD4, bind overlapping, discontinuous sites on gp120 (4,6). This complicates synthetic vaccine development since the identification and mimicry of discontinuous epitopes has proved challenging. (7)

On the other hand, both Fab12 and sCD4 share a requirement for a pair of conserved amino acids, Asp^{368} and Glu^{370} (gp120 HXBc2), in the third constant region (C3) of gp120 (4,6). While attempts to stimulate antibodies that bind gp120 with linear peptides from this region have failed (8), linear peptides are largely disordered in water and poor mimics of the ordered surfaces of native proteins. The binding pockets of antibodies to disordered peptides reflect this disorder and bind poorly if at all to the ordered surfaces of proteins. Since constrained peptides can better mimic the conformations of neutralizing epitopes (9) and stimulate higher affinity antibodies (10), we have initiated studies on the immunogenicity of a constrained C3 peptide, [JKQSSGG**D**P**E**IVGZ]C-NH_2, that includes Asp^{368} and Glu^{370} (in bold). Polyclonal antibodies raised against the constrained C3 peptide but not the linear peptide bind baculovirus recombinant gp120 IIIB.

Materials and Methods

Materials: Ethyl hydrazinoacetate hydrochloride and hydroxybenzotriazole were purchased from Aldrich. Fmoc-chloride, PyBOP, Rink amide resin were purchased from Novabiochem. Protected amino acids were purchased from Bachem, CA. Solvents are Burdick and Jackson solvents purchased from VWR.

CD-1 mice were raised at The Scripps Research Institute. Ribi's adjuvant was from Ribi ImmunoChem Research. High binding microtiter plates were from Costar. Modified Dulbecco's phosphate buffered saline (PBS), maleimide activated BSA, BSA and alkaline phosphatase-conjugated goat anti-mouse IgG $F(ab')_2$ were from Pierce Chemical Co. p-Nitrophenyl phosphate substrate, Tween 20, diethanolamine, $MgCl_2$, and NaN_3 were from Sigma. HIV-1 gp120 IIIB was from Intracel.

Sensor chip CM5 and surfactant P20 were purchased from Pharmacia Biosensor. N-hydroxysuccinimide and N-ethyl-N'-(3-diethylaminopropyl)carbodiimide were from Pierce Chemical Company. Ethanolamine and EDTA were from Sigma. Hepes, NaCl, NaOH, sodium acetate and phosphoric acid were from Fisher.

Abbreviations. Amino acids are indicated by the single letter code. Side chain protecting groups are placed in brackets. (Acn) refers to acetone that protects the terminal amino group of hydrazino acetate by forming a Schiff base. Room temperature is rt, hour(s) is h, NMP is N-methyl pyrrolidone, DIEA is diisopropyl ethylamine, DMF is dimethylformamide, DCM is dichloromethane, HOBT is hydroxybenzotriazole, TFE is trifluoroethanol, TFA is trifluroacetic acid; DIC is diisopropyl carbodiimide, EDT is ethanedithiol, d.i. H_2O is deionized water, IPA is isopropanol, DMS is dimethyl sulfide; EDC is N-ethyl-N'-(3-diethylaminopropyl)carbodiimide, FAB MS is fast atom bombardment mass spectroscopy and NMR is nuclear magnetic resonance.

FmocZ(Acn). Ethyl hydrazinoacetate hydrochloride (15.4 g, 100 mmol) was dissolved in 300 ml of a 50% aqueous solution of acetone, refluxed for 10 min, and cooled to rt. The ethyl ester was then hydrolyzed by adding NaOH (8.4 g, 210 mmol) and stirring for 30 mins at rt. The mixture was treated with Na_2CO_3 (10.6 g, 100 mmol, in 50 mL H_2O) and cooled in an ice bath. Then a solution of Fmoc-chloride (25.9 g, 100 mmol) in 50 mL of dioxane was added dropwise for a period of 1 h. The reaction mixture was stirred overnight at rt. The product was precipitated from the mixture by acidifying with

concentrated HCl. The solid was filtered and redissolved in 100 mL of acetone, refluxed for 10 mins, and the solvent removed under vacuum. Crystallization from ethyl acetate/hexane yielded 31.4 gm (80.7%) of white crystals; mp 146-148 °C. The product was confirmed by NMR spectroscopy and FAB MS.

J was prepared by first synthesizing 5,5-dimethoxy-1-oxopentanoic acid methyl ester according to the method of Stevens et al. (11). The acetal ester (1.03 g, 5.8 mmol) was mixed with methanol (5 mL and hydrolyzed by adding NaOH (0.32 g, 8 mmol) in water (5 mL) and stirring for 3 h. The solution was rotary evaporated to remove methanol, ethyl acetate (20 mL) was added and stirred in an ice-cool bath. The mixture was acidified to pH 3 with 1 N HCl. Immediately, the organic layer was separated and the aqueous layer extracted again with ethyl acetate. Ethyl acetate solution was extracted with brine and then dried with $MgSO_4$. Evaporation of solvent yielded 0.85 g (89%) of a colorless oil. The product was confirmed by NMR spectroscopy and FAB MS.

Fmoc-GZ(Acn)C(Trt)-Rink Resin. Fmoc-GZ(Acn)C(Trt)-Rink Resin was synthesized in polypropylene mesh packets (12) using an Fmoc solid phase peptide synthesis protocol. FmocC(Trt) and Fmoc-Z(Acn) were coupled using 3 equiv. of amino acid, HOBT (3 equiv.), PyBOB (3 equiv.) and DIEA (6 equiv.) in NMP. Fmoc-G was coupled using Fmoc-glycine-chloride (3 equiv.) and DIEA (3 equiv.) in NMP. Fmoc-glycine chloride was prepared according to Carpino *et al.* (13).

Acetyl-GKQSSGGDPEIVTHSGGC-NH_2 AcGK(tBoc)QS(tBut)S(tBut)GGD(tBut) PE(tBut)IVT(tBut)H(Trt)S(tBut)GGC(Trt)-Rink resin was synthesized on Rink amide resin (35 mg, 0.78 mmol/g) with an ACT350 multiple peptide synthesizer (Advanced Chem Tech). Coupling reactions were carried out with Fmoc-amino acid (5 equiv.), HOBT (5 equiv.), DIC (5 equiv.) in NMP for 45 min. After removing the final Fmoc group, the peptide was acetylated with 15% acetic anhydride in NMP. The peptide was cleaved from the resin with mixture K (0.75 g. phenol, 0.25 mL EDT, 0.5 ml thioanisole 0.5 mL d.i. H_2O, 10 mL TFA) for 1.5 h. The crude peptide was precipitated from the cleavage solution with cold ether, centrifuged, and washed three times with ether. Lyophilization yielded 30 mg of crude peptide. HPLC purification on a Vydac C18 prep-column (2.2 x 25 cm) with a 0-60% CH_3CN gradient over 40 min gave 13.3 mg. of pure peptide [FAB-MS: 1795 (M+H)].

Cyclic[JKQSSGGDPEIVGZ]C-NH_2. Fmoc-K(tBoc)QS(tBut)S(tBut)GGD(tBut)P E(tBut)IVGZ(Acn)C(Trt)-Rink resin was synthesized by extending Fmoc-GZ(Acn)C(Trt)-Rink resin with the ACT350 synthesizer using standard Fmoc chemistry described above.The peptide-resin was then packed into a polypropylene mesh packet (12) to couple J (3 equiv) with HOBT (3 equiv), PyBOB (3 equiv), and DIEA (6 equiv.). Protected peptide was cyclized on the resin with HCl (2 equiv.) in 20% TFE/DCM (45 mL for 100 mg. resin) for 20 min. The cyclic peptide was cleaved from the resin by two 15 min treatments with 10% TFA/DCM. The crude cyclic peptide solution was rotary evaporated and the residue treated with 95% TFA/H_2O for 1 h. It was precipitated from TFA with cold ether, centrifuged and washed twice with cold ether. Lyophilization yielded 51 mg of crude peptide from 100 mg of resin. Purification on HPLC using a Vydac C18 semiprep-column with a 0-60% CH_3CN gradient over 40 min gave 1.5 mg of pure cyclic peptide. 1D NMR spectroscopy confirmed the presence of the hydrazone link. FAB mass spectroscopy confirmed the predicted mass [FAB-MS: 1248 (M+H)].

(Fmoc-G)$_2$KA-Rink Resin. Fmoc-amide-Rink resin (1.5 g, 0.45 mmol, 0.3 mmol/g) was deprotected with 30% piperidine/DMF, washed 4 times with DMF, once with IPA, and once with NMP. Fmoc-A (47 mg, 0.15 mmol) was coupled using HOBT (20.2 mg, 0.15 mmol), PyBOB (78 mg, 0.15 mmol) and DIEA (0.052 ml, 0.3 mmol) in NMP for 40 min. The resin was then washed with DMF and DCM and treated with 15% acetic anhydride/NMP for 30 min to cap free amino groups. Fmoc-K(Fmoc) and Fmoc-G were coupled as describe above.

MAPS core ($ClCH_2CO$-QYIKANSKFIGITELKKK)$_4$(KGG)$_2$KA-NH_2.(FmocQY(tBut)IK(tBoc)ANS(tBut)K(tBoc)FIGIT(tBut)E(tBut)LK(tBoc)K(tBoc)K(tBoc))$_4$(KGG)$_2$KA-Rink resin was synthesized with the ACT350 multiple peptide synthesizer using (Fmoc-G)$_2$KA-Rink resin as starting material. The final Fmoc protecting group was removed with 30% piperidine in DMF and the resin washed with DMF, IPA and NMP. Peptide-Rink resin (50 mg, 0.0005 mmol) was treated with $(ClCH_2CO)_2O$ (68.4 mg, 0.4 mmol) in DMF (10 mL) for 40 min and then washed with DMF and DCM. The peptide was cleaved from the resin by treatment with 5 mL solution of TFA:DMS:anisole (96:2:2) for 1 h. Precipitation with ether and lyophilization yielded 9 mg of crude peptide. Purification on an analytical Vydac C4 column with a gradient of 0-10-35% acetonitrile over 0-5-35 min yielded 1.1 mg of pure MAPS core [electrospray-MS: 9,373 (MH+)].

Loop-MAPS. Chemical ligation was carried out using Tam's strategy (14). Solvents for this reaction were degased and purged with nitrogen. Cyclic[JKQSSGGDPEIVGZ]C-NH_2 (0.6 mg, 0.48 mmol) and ($ClCH_2CO$-QYIKANSKFIGITELKKK)$_4$(KGG)$_2$KA-NH_2 (0.3 mg, 0.032 mmol) were dissolved in 6M guanidinium hydrochloride solution (0.2 mL), 1M Tris HCl solution (0.2 mL) added, and the reaction mixture stirred under N2 Progress was followed by HPLC on an analytical Vydac C4 column. The product was purified by HPLC on the same column yielding 0.27 mg of 90% pure four branched Loop-MAPS [electrospray-MS: 14,937 (MH+)]. A minor amount of two-branched and three-branched MAPS was observed by polyacrylamide gel electrophoresis (Fig. 4).

Linear-MAPS. Linear-MAPS was synthesized using the protocol described in (7), Ac-GKQSSGGDPEIVTHS-GGC-NH_2 (1.5 mg, 0.83 mmol), and (ClCH2CO-QYIKANSKFIGITELKKK)4(KGG)2KA-NH_2 (0.5 mg, 0.05 mmol) yielded 0.45 mg of pure four-branched Linear-MAPS [electrospray-MS: 16,255 (M+H+)].

Immunization. Three CD-1 outbred mice (M1-M3) were immunized by intraperitoneal injection with 50 ug Loop-MAPS in complete Freund's adjuvant (Sigma) and boosted three times with subcutaneous injections at three week intervals with the same quantity of MAPS in the Ribi adjuvant system R-700. The process was repeated with Linear-MAPS in M3-M5. The mice were prebled (Bleed 1) before the first immunization and bled again immediately before the second immunization (Bleed 2), one week later (Bleed 3), at the third immunization (Bleed 4) and every week after for four weeks (Bleeds 5-8), at the fourth immunization (Bleed 9) and every week thereafter for three weeks (Bleeds 10-12). One of the mice died a day within a day of the fourth injection with the Linear-MAPS.

ELISAs. Antiserum from each of the mouse sera was titered using enzyme-linked immunoadsorbent assays (ELISAs) with linear and loop peptides conjugated to maleimide activated BSA (Pierce Chemical Co.) as well as baculovirus recombinant gp120 IIIB (Intracel). Peptides were conjugated to maleimide-BSA according to the manufacturer. ELISA plates were coated overnight at 4 C with peptide conjugate (0.2 ug/well in 50 uL of PBS) or gp120 IIIB (0.1 ug/well in 50 uL of PBS), the plate was washed four times with H_2O, blocked with 50 uL of PBS containing 1% BSA and 0.05% Tween 20 for 1 h at rt and incubated overnight at 4 C with 50 uL of mice sera diluted with PBS containing 1% BSA and 0.05% Tween 20. After 10 washes with water, 25 uL of 1:500 dilution of alkaline phosphatase-conjugated goat antimouse IgG-F(ab')$_2$ was added and incubated at 37 C for 1 h. This was followed with 10 washes of water before adding 50 uL of

developing solution (5 mg of p-nitrophenyl phosphate in 5 ml of developing buffer: 50 mL of diethanolamine, 50 mg of $MgCl_2$ and 97.5 mg of NaN_3 in 500 mL water, pH 9.7). Color development was recorded at 405 nm after 15 min with a Bio-Tek microplate reader.

Surface plasmon resonance. Immobilization of HIV-1 gp120 to the sensor chip surface was performed according to the amine coupling procedure recommended by the manufacturer (Pharmacia). The immobilization protocol was preformed with a continuous flow of HBS of 5 ml/min. Briefly, after equilibration of the instrument with HBS buffer (10 mM Hepes, 150 mM NaCl, 0.005% surfactant P20, pH 7.4), equal volumes of 0.05 M NHS and 0.2 M EDC were machine mixed and 35 uL injected across the surface to activate the carboxyl groups. Then gp120 IIIB (40 ul, 20 mg/ml in 10 mM sodium acetate, pH 4.7) was injected across the activated surface, residual NHS-ester capped with 1M ethanolamine, pH 8.5 (15ul) and the surface regenerated by removing noncovalently bound gp120 IIIB with 1M phosphoric acid (15 ul). Similar results have been reported by Brigham-Burke, M. et al. (15)

Binding of anti-peptide polyclonal sera to the immobilized gp120 IIIB. Each binding cycle was performed with a constant flow of 5 ml/min of HBS. The sera were diluted with HBS (1/50), filtered through a 0.45 mm cellulose membrane and injected (50 uL) across the immobilized gp120. The surface was regenerated by injection of 0.1 M phosphoric acid (15 uL).

Competition of anti-peptide polyclonal sera with peptides for gp120. Mouse 3 (M3) antiserum (1/50 in HBS) was incubated with Loop or Linear peptides 30 min before injection (50 uL) across the immobilized gp120. The surface was regenerated with an injection of 0.1M phosphoric acid (15 uL).

Results

Peptide Design. Olshevsky et al. (4) proposed that the conserved Asp^{368}-Pro-Glu^{370} sequence in the C3 region on gp120 (HXBc2) formed an exposed beta turn. GORBturn (16), a secondary structure algorithm for turn types, predicts a strand-(three overlapping, nonspecific turns)-strand structure for the C3 sequence, FKQ<u>SSGG**D**P**E**</u>IVT. Asp^{368} and Glu^{370} are in bold and the predicted turn region is underlined. This prediction is consistent with the exposure of a hydrophilic β-hairpin loop on the viral surface (Fig. 1). β-hairpin loops are characterized by a ladder of hydrogen bonds that span alternate pairs of amino

acids (Fig. 1). However, hydrogen bonds are weak and insufficient for stabilizing peptide conformations in water. We replaced one of the predicted hydrogen bonds with a hydrazone covalent hydrogen bond mimic using modified amino acids as outlined in Figure 2 (17). The expectation was that if the prediction was correct that the mimic would stabilize the peptide in a native conformation (9). The constrained peptide could then be used to stimulate antibodies with binding pockets that would bind complementary structures on native gp120 (10).

MAPS design and synthesis. In order to stimulate a polyclonal antibody response to the C3 peptide, it should be linked to a T-cell epitope. Multiple antigen presentation systems (MAPS) provide a means for linking B and T-cell epitopes in chemically defined systems that stimulate strong immune responses (18). Both linear and loop peptides ending in C-terminal cysteines were chemically ligated to synthetic chloroacetylated "universal" tetanus toxoid epitope on a four-branch MAPS core (Fig. 3). The products of this reaction, a Linear-MAPS and a Loop-MAPS, were purified on an analytical Vydac C4 column. The composition was confirmed by high density gel electrophoresis and electrospray mass spectroscopy (Fig. 4). These small proteins (15 kD) are completely soluble and provide well characterized vehicles for "universal" T-cell epitopes and constrained peptides.

Both Linear and Loop-MAPS were immunogenic in each of the outbred mice tested indicating that T-cell help can be stimulated by these fully synthetic vaccines (Fig. 5). Five out of the six mice show high titers against both linear and loop peptides. All but one of the antisera show higher titers against the loop peptide than the linear peptide. This establishes that antibodies formed in mice immunized with the linear peptide do not require the free end of a peptide for binding. Despite this capability, antisera formed in response to the linear peptide did not bind recombinant gp120 IIIB (Fig. 5). On the other hand, two of the three mice (M1,M3) immunized with the Loop-MAPS show titers to gp120 IIIB (Fig. 5) indicating the epitope is exposed. This suggests that the antibodies formed in response to the loop peptide have special properties that allow it to bind to the epitope. Titers for the best responders to the Linear-MAPS (M5) and Loop-MAPS (M3) for bleeds from the first three immunizations confirm the different responses to gp120 (Fig. 6).

The ELISA titer for M3 antiserum indicating that antibodies formed in response to the Loop-MAPS bind gp120 IIIB was confirmed by surface plasmon resonance measurements. Gp120 was first immobilized on the sensor surface (Fig. 7). M3 antiserum binds to immobilized gp120 (300 RU) whereas M5 antiserum formed in response to

Linear-MAPS binds poorly if at all (100 RU)(Fig 8). Competition experiments confirmed the specificity of the M3 antisera reaction with gp120. Preincuation of M3 antiserum with either 7 uM Loop or Linear peptides completely blocked binding to gp120 IIIB while 350 nM of Loop and Linear peptide blocked binding by 59% and 54% respectively (Fig. 9).

Discussion

Studies with site-directed mutants of gp120 show that gp120/CD4 interaction is dependent on Asp-Pro-Glu from the C3 region. A whole class of potent neutralizing antibodies that block gp120/CD4 binding share this dependency. The sequence is predicted to form a hydrophilic loop on the surface of gp120 and the observation that antibodies depend on an intact sequence suggest that it is exposed and/or accessible to antibody binding pockets. Alternatively, the Asp-Pro-Glu sequence could play a structural role. Substitution of Asp or Glu by other amino acids could trigger a conformational change in gp120 that would preclude binding to a discontinuous epitope defined by other amino acids.

One approach to resolving this question is to raise antibodies to the Asp-Pro-Glu sequence and test whether they can bind native gp120. McKeating et al. (8) injected three rats with a linear peptide, 730-34 (KQSSGG**DPE**IVTHSFNCGGE) conjugated to ovalbumin. While antisera bound the peptide (half-maximal binding at 1,200-1,500), no binding to gp120 was dectected. Our own studies with a *P. falciparum* malaria peptide showed that one can improve antigenicity and immunogenicity of a peptide by constraining it. Antibodies raised against a loop peptide that adapts a native conformation in water bind and agglutinate the malaria parasite while antibodies against the corresponding linear peptide do not. Thus McKeating's negative result with a linear peptide does not necessarily mean that an epitope is not exposed. However, it may mean that the peptide must be correctly folded in order to stimulate gp120 binding antibodies.

In this study, we predicted a conformation for the C3 sequence, KQSSGG**DPE**IV, and constrained the peptide with a covalent hydrogen bond mimic. This was our best guess. Our experience with the *P. falciparum* peptide (10) and a gp120 V3 peptide (19) shows that peptide size, cadence and link can effect the degree of conformational mimicry to a considerable degree. For this reason, we prefer to initially show enhanced binding of a constrained peptide to antibodies before proceeding. Enhanced binding provides a direct measure of conformational mimicry. However, sCD4 and Fab12 do not bind the loop we made for this study. This suggests that either the loop does not adapt a native like conformation or that sCD4 and Fab12 binds a more extensive epitope as is likely.

Although we can not be sure that the synthetic C3 peptide loop adapts a native conformation, we can test it as an immunogen. If it adapted a native like conformation, it could stimulate antibodies that might bind native gp120. Much to our surprise, the Loop-MAPS but not the Linear-MAPS stimulates a polyclonal response in mice that includes a fraction of antibodies that bind recombinant gp120. Since antisera to both Loop-MAPS and Linear-MAPS bind the free loop peptide in ELISAs, these antisera should be able to bind internal epitope(s) in gp120 i.e. free ends are not a requirement for binding. However, only antisera to the Loop-MAPS binds recombinant gp120. This shows that the epitope(s) exposed on the free loop peptide is accessible on gp120 to loop antiserum but not linear antiserum. These antisera thus show differences in either fine structure and/or conformational specificities that become apparent in gp120 binding reactions. This suggests that the epitope on recombinant gp120 IIIB bound by the loop antiserum is folded or buried to some degree making it inaccessible to linear peptide antisera. It is otherwise difficult to account for the difference in affinity for gp120. Thus it is likely that the Loop-MAPS presents a relevant conformation to the immune system.

References

1. La Rosa, G. J. et al., (1990). Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant, Science 249:932 .

2. Trkola, A. et al., (1995). Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG, J. Virol. 69:6609 .

3. Barbas, C. F. et al., (1992). Recombinant human Fab fragments neutralize human type 1 immunodeficiency virus *in vitro*, Proc. Natl. Acad. Sci. USA, 89:9339

4. Olshefsky, U. et al., (1990). Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4-receptor binding, J. Virol. 64: 5701

5. Ditzel, H. J. et al., (1995). Neutralizing recombinant human antibodies to a conformational V2- and CD4-binding site-sensitive epitope of HIV-1 gp120 isolated by using an epitope-masking procedure, J. Imm. 154: 893

6. Roben, P. et al., (1994). Recognition properties of a panel of human recombinant Fab fragments to the CD4 binding site of gp120 that show differing abilities to neutralize human immunodeficiency virus type 1, J. Virol. 68 (8): 48217.

7. Bonnycastle, L.L.C. et al., (1996). Probing the basis of antibody reactivity with a panel of constrained peptide libraries displayed by filamentous phage, J. Mol. Biol. 258:747

8. McKeating, J.A. et al., (1992). Monoclonal antibodies to the C4 region of human immunodeficiency virus type 1 gp120: use in topological analysis of a CD4 binding site, AIDS Res. Hum. Retroviruses 8: 451

9. Cabezas, E. et al., (1996). Identifying determinants of protein structure with loop peptides, Peptides: Chemistry, Structure and Biology (Proceedings of the Fourteenth American Peptide Symposium) P.T.P. Kaumaya and Hodges, R.S. (Eds.) Mayflower Scientific Ltd. pp. 734-735

10. Satterthwait, A. C. et al., (1996). Constrained synthetic peptide vaccines, Peptides: Chemistry, Structure and Biology (Proceedings of the Fourteenth American Peptide Symposium) P.T.P. Kaumaya and Hodges, R.S. (Eds.) Mayflower Scientific Ltd. pp. 772--773

11. Stevens et al, (1979). On the Stereochemistry of the Robinson-Schopf Reaction. A stereospecific total synthesis of the ladybug defense alkaloids precoccinelline and coccinelline, J. Am. Chem. Soc., 101:23, 7032-7035

12. Houghten, R. et al. (1986). Simultaneous multiple peptide synthesis: the rapid preparation of large numbers of discrete peptides for biological, immunological and methodological studies, Biotechniques 4(6):522

13. Carpino, et al., (1986). J. Org. Chem. : 51, 3732-3734

14. Defoort, J-P. et al., (1992). A rational design of slynthetic peptide vaccine with a built-in adjuvant, Int. J. Peptide Protein Res. 40:212

15. Brigham-Burke, M. et al, (1992). Detection of receptor-ligand interactions using surface plasmon resonance: Model studies employing the HIV-1 gp120/CD4 interaction, Anal. Biochem. 205 125-131

16. Wilmot, C.M. and Thornton, J. M. (1988). Analysis and prediction of the different types of β-turn in proteins, J. Mol. Biol. 203:221

17. Chiang, L-C. et al., (1994).The synthesis of peptide secondary structure mimetics with covalent hydrogen bond mimics on the solid support, Peptides: Chemistry, Structure and Biology (Proceedings of the Fourteenth American Peptide Symposium) Hodges, R.S. and Smith, J. A. (Eds.) ESCOM, Leiden, pp. 278-280

18. Tam, J. (1988). Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigen, Proc. Natl., Sci. USA 85:5409

19. Cabezas, E. and Satterthwait, A. C., (1996). Defining conformational requirements for the principle neutralizing determinant on HIV-1, Peptides: Chemistry, Structure and Biology (Proceedings of the Fourteenth American Peptide Symposium) P.T.P. Kaumaya and Hodges, R.S. (Eds.) Mayflower Scientific Ltd, pp. 800-801

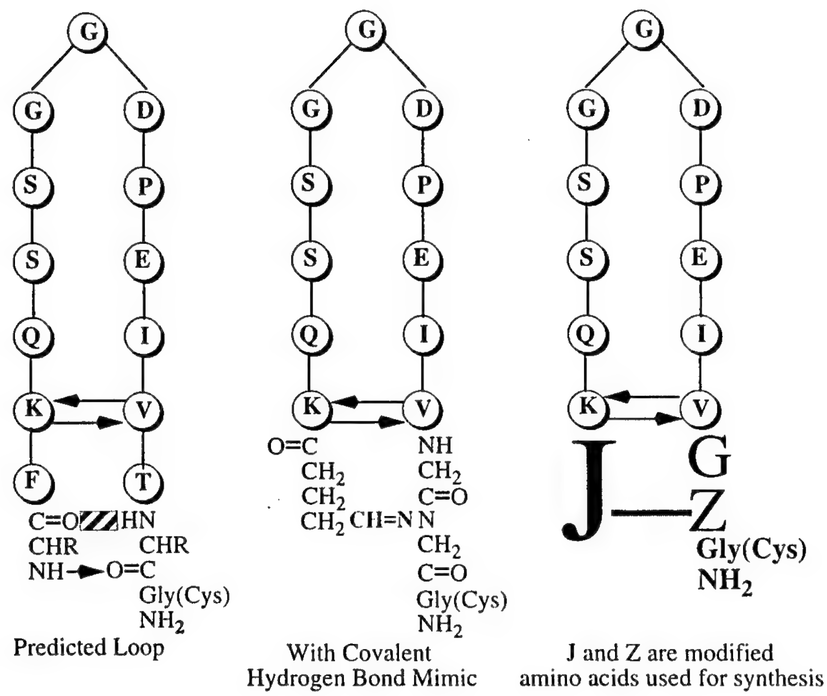


Figure 1. Predicted and designed loop peptide with a hydrogen bond mimic.

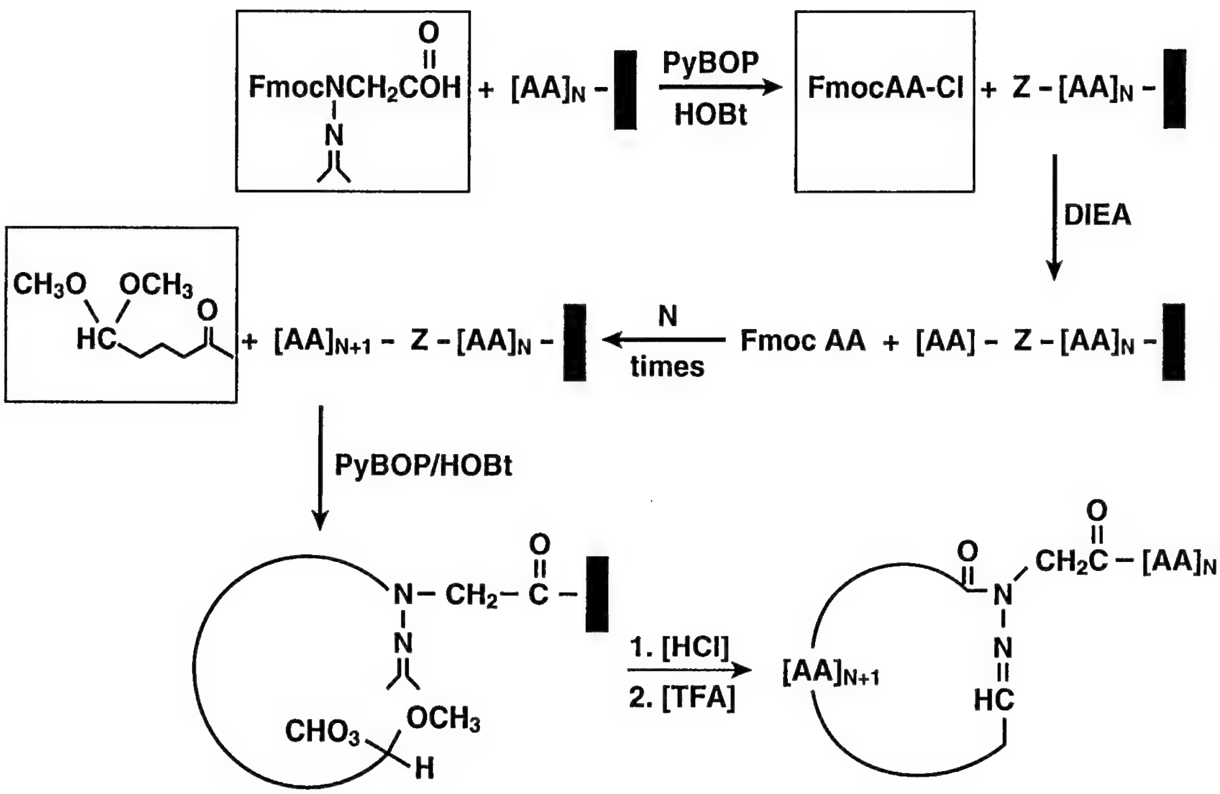
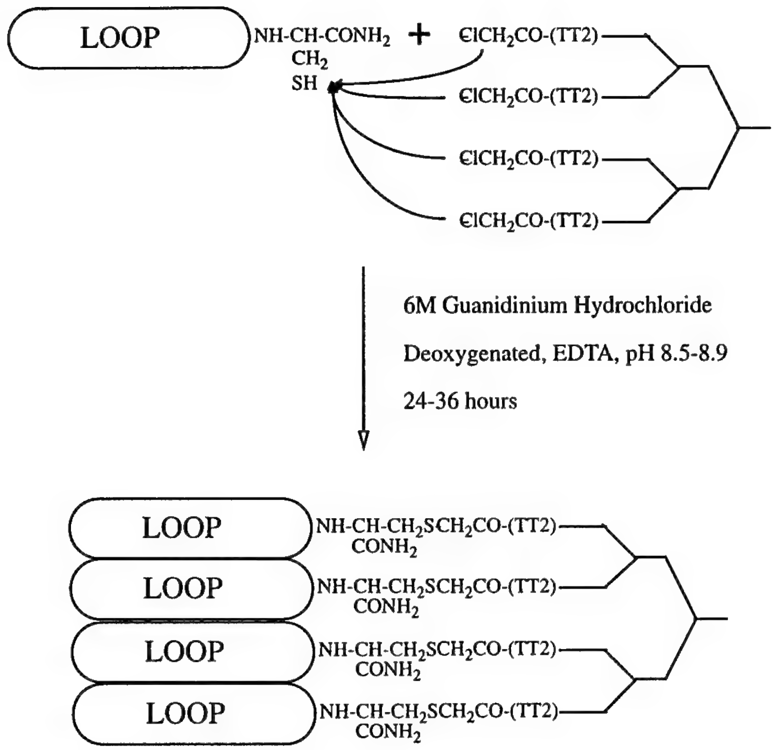


Figure 2. Synthesis of a loop peptide with a hydrazone covalent hydrogen bond mimic by solid phase peptide synthesis. Key intermediates are boxed. The syntheses of these intermediates and a loop peptide are described in detail in Materials and Methods.



TT2: QYIKANSKFIGITEL

Linear: AcGKQSSGGDPEIVTHSGGC-NH_2

Loop: JKQSSGGDPEIVGZC-NH_2

Figure 3. Chemical ligation of linear and loop C3 peptides to a "universal" tetanus toxoid T-cell epitope (TT2) to yield a four-branched multiple antigen presentation system (MAPS).

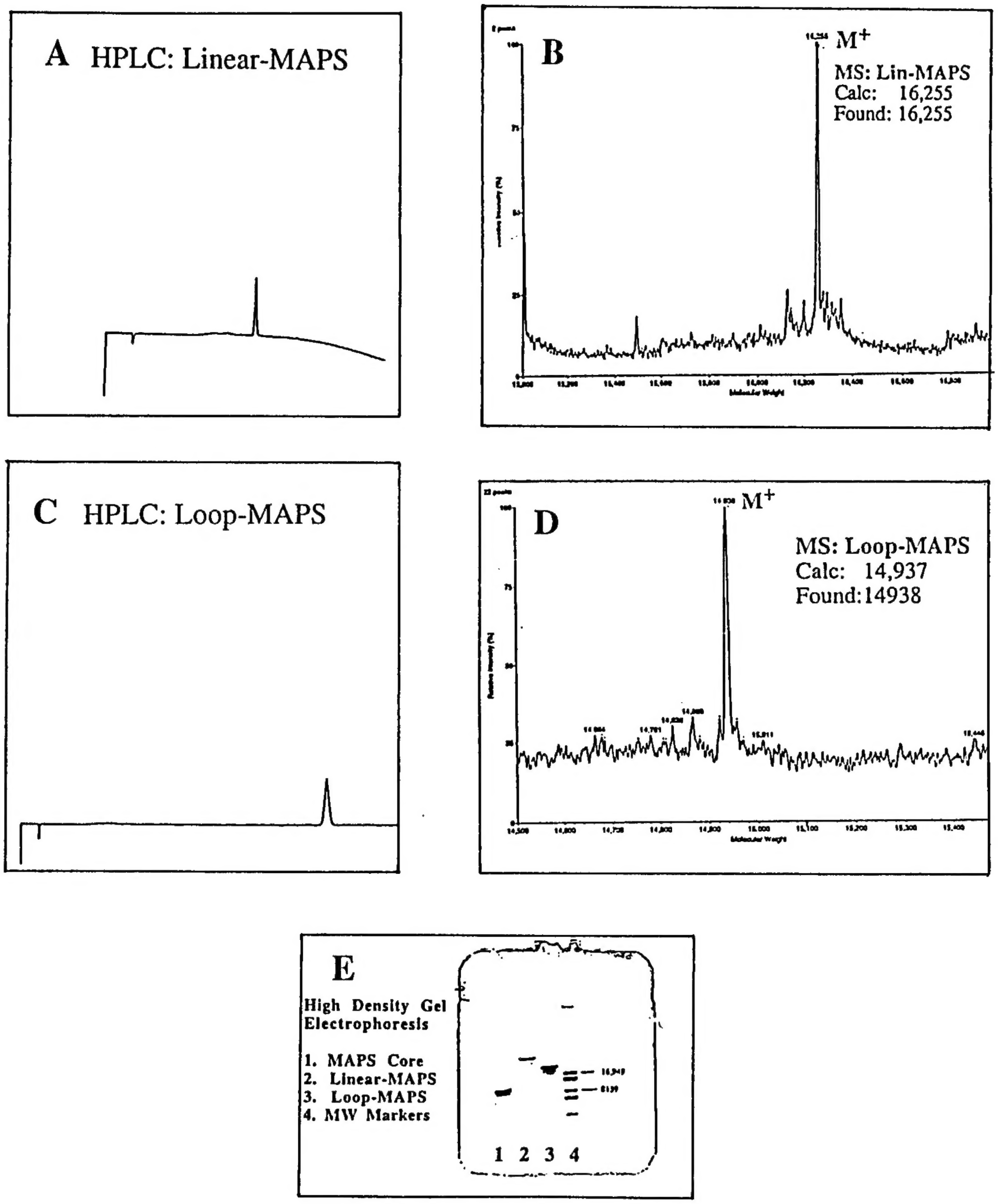


Figure 4. Soluble Linear and Loop-MAPS were synthesized in nearly pure form as demonstrated by high pressure liquid chromatography, electrospray mass spectroscopy high density gel electrophoresis.

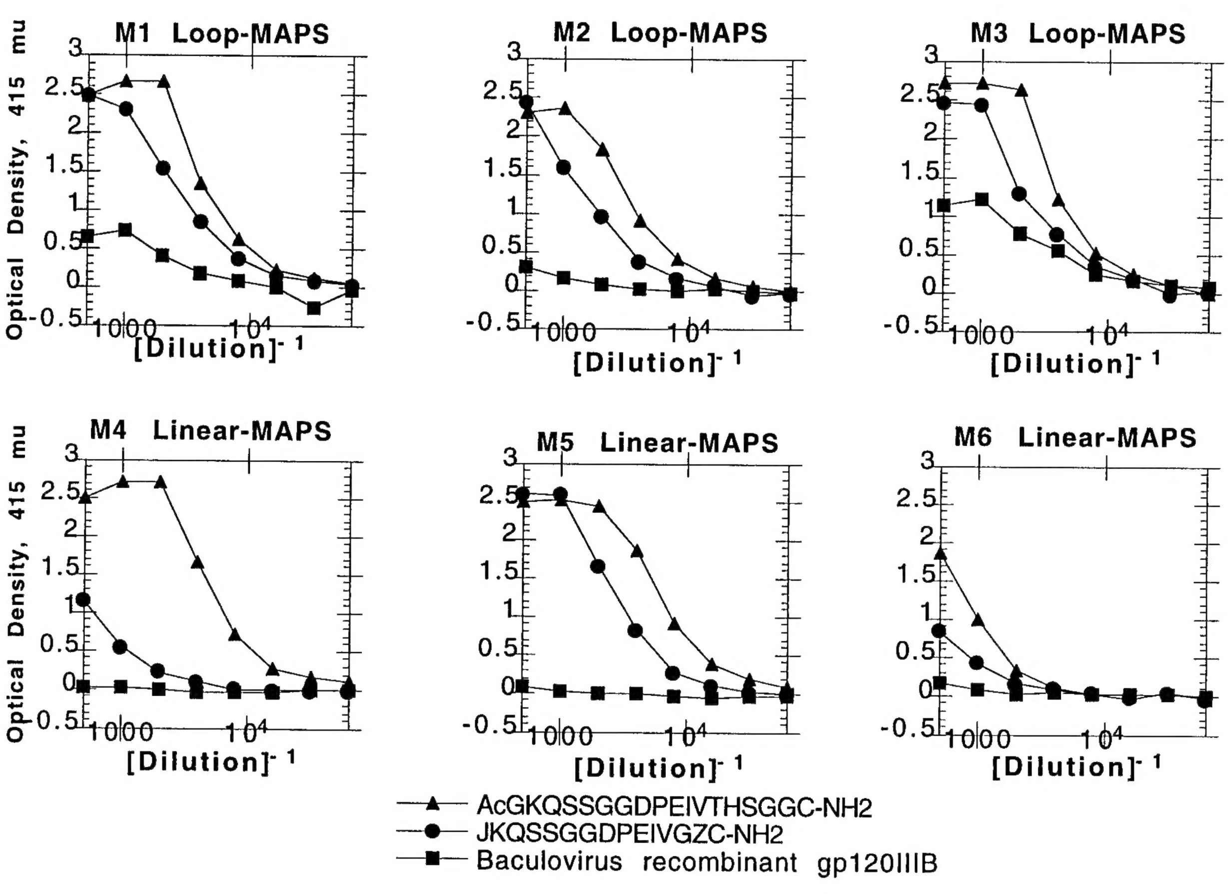


Figure 5. ELISA titers for mice immunized with Loop-MAPS (fifth bleed, M1,M2.M3) and Linear-MAPS (fourth bleed, M4,M5,M6). Antisera were titered against three antigens: Linear peptide, Ac-GKQSSGGDPEIVGGC-NH_2 (triangles), Loop peptide cyclic-[JKQSSGGDPEIVGZ]C-NH_2 (circles) and baculovirus recombinant gp120 IIIB (squares). Only antisera from mice immunized with Loop-MAPS binds recombinant gp120 IIIB.

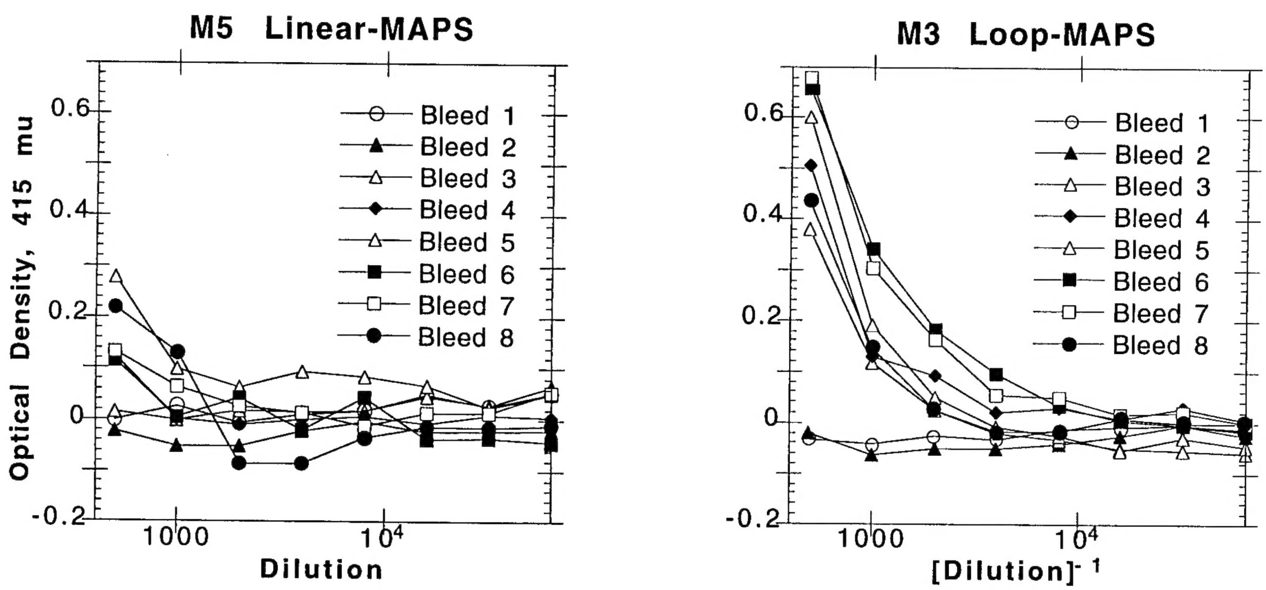


Figure 6. ELISA titers of antiserum from Mouse 5 (left panel) injected with Linear-MAPS and Mouse 3 (right panel) injected with Loop-MAPS to recombinant gp120 IIIB. Antisera are from Bleed 1 (preimmunization), Bleed 2 (before 2nd immunization at 3 weeks), Bleed 3 (1 week after the third immunization), Bleed 4 (before 3rd immunization at 6 weeks), Bleeds 5-8 (at 1,2,3,4 weeks following the 3rd immunization).

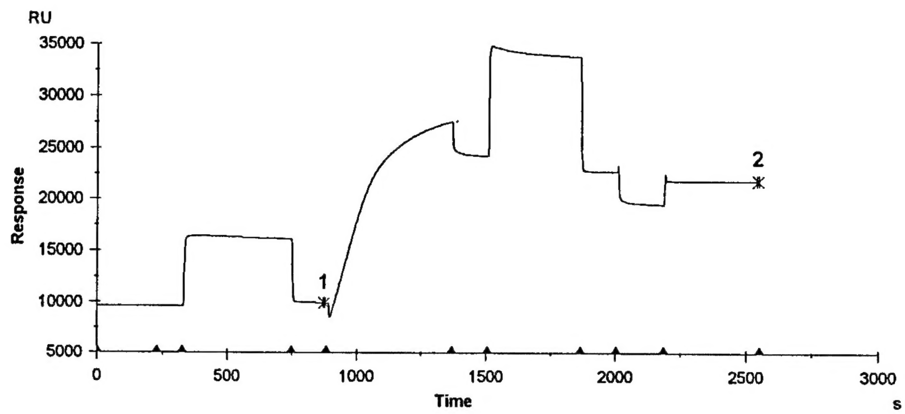


Figure 7. Immobilization of gp120 IIIB on the biosensor surface. The difference between report points 1 and 2 (11,818 RU) corresponds to the immobilized gp120.

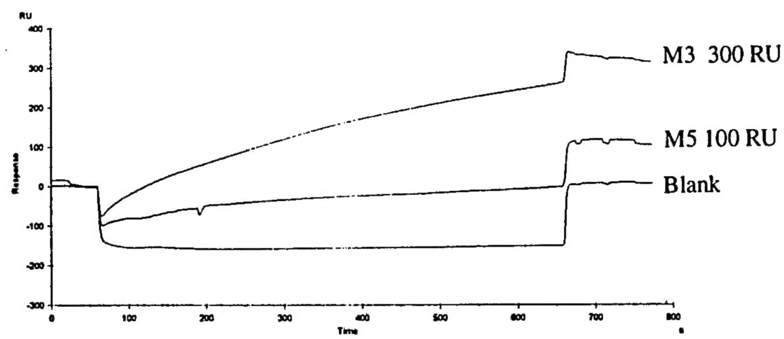


Figure 8. Sensorgram obtained for injecting antiserum (50 ul, 1/50 dilution in HBS buffer) from Mouse 3 (anti Loop-MAPS) and Mouse 5 (anti Linear-MAPS) over immobilized gp120 IIIB. The flow rate is 5 ul/min. The blank measures response to a 50 ul injection of HBS buffer.

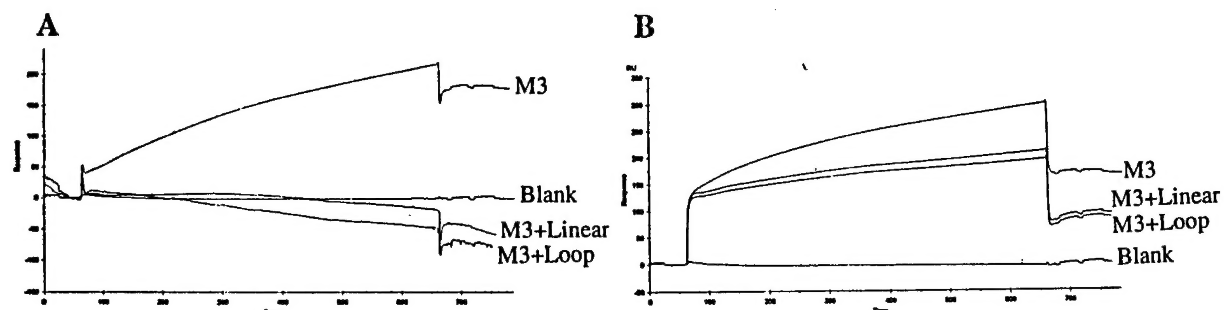


Figure 9. Sensorgram showing the inhibition of binding of Mouse 3 antiserum (50 ul 1/50 dilution in HBS buffer) to immobilized gp120 IIIB with (A) 7 μm linear or loop peptide and (B) 350 nM peptide. The flow rate is 5 ul/min. The blank measures response to a 50 ul injection of HBS buffer.